MATHEMATICAL MONOGRAPHS.

EDITED BY

MANSFIELD MERRIMAN AND ROBERT S. WOODWARD.

No. 12.

THE THEORY OF RELATIVITY

BY

ROBERT D. CARMICHAEL,

PROFESSOR OF MATHEMATICS IN UNIVERSITY OF ILLINOIS

SECOND EDITION

NEW YORK

JOHN WILEY & SONS, INC.

LONDON: CHAPMAN & HALL, LIMITED

1920

PRESS OF
BRAUNWORTH & CO.
BOOK MANUFACTURERS
BROOKLYN, N. Y.

PREFACE

THE theory of relativity has now reached its furthest conceivable generalization in the direction of the covariance of the laws of nature under transformations of coordinates. The older theory of relativity remains valid as a special case of the general theory and may well serve as an introduction to its more far-reaching aspects. Accordingly, in the present (second) edition of this monograph I have retained the older theory in precisely the same form as in the first edition, the matter covering Chapters I to VI of the present treatment, and have added the longer Chapter VII to give a compact account of the generalized theory. The tendency now is to call the latter the theory of relativity and to distinguish the older from it by giving to the older theory the name of the restricted theory of relativity.

In the opening section (§ 37) of the new chapter, I give a brief summary of results from the restricted theory. Anyone who is acquainted with these, whether derived as in this book or otherwise, may proceed at once to the reading of Chapter VII. It is believed that he will find in it about as brief an account of the new theory as can be given so as to be easily intelligible and at the same time to reach the general theory of gravitation, to make clear the nature of the three famous crucial phenomena, to associate the theory with Maxwell's electromagnetic equations, and to place the whole in its proper setting with respect to the general body of scientific truth.

The new as well as the older matter in the booklet has been written from the point of view of the usefulness of the theory of relativity in the development of physical science. No applications are given other than those which are directly and immediately associated either with the fundamental ideas or

with certain crucial phenomena for testing the validity of the theory. In this way only may the central elements of novelty most readily be brought to light.

No attempt has been made to give a complete account of the theory. The purpose of the monograph is best served by presenting only those fundamental developments which are needed for and contribute directly to making clear the main characteristics of the theory. The more detailed statements are to be found elsewhere, especially in the memoirs which have now reached a considerable number.

Every exposition of the general theory of relativity must be deeply indebted to the basic memoir of Einstein, published in 1916 in *Annalen der Physik*, volume 49. Very useful to me also, as every reader will observe, has been the report of A. S. Eddington to the Physical Society of London on "The Relativity Theory of Gravitation," a booklet to which one may be referred who wishes to go further into the theory than the exposition of the present monograph will carry him.

R. D. CARMICHAEL.

UNIVERSITY OF ILLINOIS,
April, 1920.

CONTENTS.

CHAPTER I. INTRODUCTION.

CHAPTER II. THE POSTULATES OF RELATIVITY.

CHAPTER III. THE MEASUREMENT OF LENGTH AND TIME.

CHAPTER IV. EQUATIONS OF TRANSFORMATION.

CHAPTER V. MASS AND ENERGY.

CHAPTER VI. EXPERIMENTAL VERIFICATION OF THE THEORY.

CHAPTER VII. THE GENERALIZED THEORY OF RELATIVITY.

THE THEORY OF RELATIVITY.

CHAPTER I.

INTRODUCTION.

§1. The Foundations of Physics.

Those who look on physics from the outside not infrequently have the feeling that it has forgotten some of its philosophical foundations. Even among its own workers this condition of the science has not entirely escaped notice.

The physicist, who, above all other men, has to deal with space and time, has fallen into certain conventions concerning them of which he is often not aware. It may be true that these conventions are just the ones which he should make. It is certain, however, that they should be made only by one who is fully conscious of their nature as conventions and does not look upon them as fixed realities beyond the power of the investigator to modify.

Likewise, a question arises as to what element of convention is involved in our usual conceptions of mass, energy, etc.; that the question is not easily answered becomes apparent on reflection.

These and many other considerations suggest the desirability of a fresh analysis of the foundations of physical science. Now it is a ground of gratulation for all those interested in this matter that there has arisen within modern physics itself a new movement—that associated with the Theory of Relativity—which is capable of contributing most effectively to the

construction of a more satisfactory foundation for its superstructure of theory.

It is at once admitted that the theory of relativity is not yet established on an experimental basis which is satisfactory to all persons; in fact, some of those who dispute its claim to acceptance are among the most eminent men of science of the present time. On the other hand there is an effective body of workers who are pushing forward investigations the inspiration for which is afforded by the theory of relativity.

This state of affairs will probably give rise to a considerable controversial literature. If the outcome of this controversy is the acceptance in the main of the theory of relativity, then this theory will afford just the means needed to arouse in investigators in the field of physics a lively sense of the philosophical foundations of their science. If the conclusions of relativity are refuted this will probably be done by a careful study of the foundations of physical science and a penetrating analysis of the grounds of our confidence in the conclusions which it reaches. This of itself will be sufficient to correct the present tendency to forget the philosophical basis of the science.

It follows that in any event the theory of relativity will force a fresh study of the foundations of physical theory. If it accomplishes no more than this it will have done well.

§ 2. Are the Laws of Nature Relative to the Observer?

The fundamental question asked in the theory of relativity is this: In what respect are our enunciated laws of nature relative to us who investigate them and to the earth which serves us as a system of reference? How would they be modified, for instance, by a change in the velocity of the earth?

To put the matter more precisely, let us suppose that we have two relatively moving platforms with an observer on each of them. Suppose further that each observer considers a system of reference, say cartesian axes, fixed to his platform, and expresses the laws of nature, as he determines them, by means

of mathematical equations involving the cartesian coordinates as variables. To what extent will the laws in the two cases be identical? What transformations of the time and space variables must be carried out in order to go from the equations in one system to those in the other; that is, what relations must exist between the variables on the two platforms in order that the results of observation in the two cases shall be consistent? Any theory which states these relations is a theory of relativity.

It is obvious that the questions above must be fundamental in any system of mechanics. In fact, a detailed analysis of the matter would show that such a system is characterized primarily by the answers which it gives to these questions. This is the feature which distinguishes between the Newtonian and the various systems of non-Newtonian Mechanics. The theory of relativity, in the sense of this book, belongs to one of the latter. It is developed from a small number of fundamental postulates, or laws, which have been enunciated as the probable teaching of experiment. Some account of these experimental investigations will now be given.

§ 3. The State of the Ether.

Those who postulate the existence of an ether as a means of explaining the facts about light, electricity and magnetism have usually been in general agreement as to the conclusion that the parts of this ether have no relative motion among themselves, that is, that the ether may be considered stationary. Experimental facts, which have to be accounted for, cannot be explained satisfactorily on the hypothesis of a mobile ether.

The aberration of light is one of the most conspicuous of those phenomena which seem to require for their explanation the hypothesis of a stationary ether.

The experiment of Fizeau, in which a comparison was made between the velocities of light when going with, and against, a stream of water, was interpreted by Fresnel as indicating

a certain entrainment of the ether; but a later examination of the matter by Lorentz * has led to the conclusion that Fizeau's experiment requires a stationary ether for its explanation.

A result which leads to a similar conclusion has been obtained in electrodynamics by H. A. Wilson † in measuring the electric force produced by moving an insulator in a magnetic field.

§ 4. Movement of the Earth through the Ether.

The theory of a stationary ether leads us to expect certain modifications in the phenomena of light and electricity when there is no relative motion of material bodies, but when both the observer and all his apparatus are carried along through the ether with a velocity v. The effects to be expected are of the order v^2/c^2, where c is the velocity of light. Although these effects are very small even when v is the velocity of the earth in its orbit, the possible accuracy of certain optical and electrical experiments is such that these effects could certainly be found if they existed without some compensating effect to mask them. Thus it should be possible for an observer, by making optical and electrical measurements on the earth alone, to detect the motion of the earth relative to the ether.

§ 5. The Test of Michelson and Morley.

Thus it was predicted that the time which would be required for a beam of light to pass a given distance and return would be different in the two cases when the path of light was parallel to the direction of motion and when it was perpendicular to this direction. Michelson and Morley ‡ devised an experiment the object of which was to put this prediction to a crucial test.

The experiment was a bold one, seeing that the difference to be measured was so small; but it was carried out in such a

* See Lorentz, Versuch einer Theorie der Elektrischen und Optischen Erscheinungen, in Bewegten Körpern, § 68.

† Proc. Roy. Soc. 73 (1904): 490.

‡ *American Journal of Science* (3), 34 (1887): 333–345.

brilliant way as to permit no serious doubt of the accuracy of the results. The difference of time predicted by theory was found by experiment not to exist; there was not the slightest difference of time in the passage of light along two paths of equal length, one in a direction parallel to the earth's motion and the other in a direction perpendicular to it.

Owing to the great importance which this famous experiment has in the theory of relativity some further account of it will be given here. The essential parts of the apparatus used are shown in Fig. 1, and the experiment was carried out in the following manner:

Let a ray of light from a point source S fall on a semi-reflecting mirror A, which is set at such an angle that it will reflect

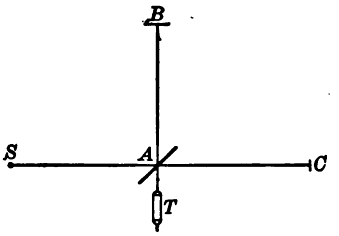

FIG. 1.

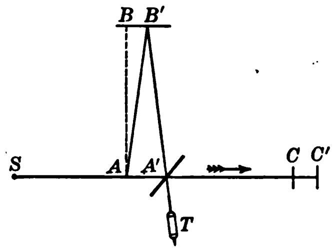

FIG. 2.

half the ray to the mirror B and allow the other half to pass on to a third mirror C. The lines AB and AC cross at right angles and the distance AB is made equal to the distance AC. Half of the reflected ray from B will pass through A and on to the telescope T. Also, half of the reflected ray from C will be reflected at A to T. Now the paths $ABAT$ and $ACAT$ are by measurement equal, so that the ray along ABA and the one along ACA should reach T simultaneously, provided that the apparatus is at rest in the ether.

Now suppose that the ether is stationary and that the earth is moving through it with little or no disturbance. Then the whole system of apparatus, which is fixed to the earth, will be moving with respect to the ether with the effect indicated in Fig. 2.

While the light is going from mirror A to mirrors B and C and back again to A, the whole apparatus is carried forward in the direction of the incident light to the position $A'B'C'$. The ray reflected from B, which interferes with a given ray from C along the line $A'T$, must be considered as traveling along the line $AB'A'$, the angle BAB' being the angle of aberration.

Suppose that the ether remains at rest. Denote by c the velocity of light, and by v the velocity of the apparatus. Let t be the time required for the light to pass from mirror A to mirror C, and let t' be the time required in returning from C to A'. At the time when the reflection takes place at the mirror C, this mirror is approximately half way between C and C' of the figure.

Let D represent the distance AB or AC. Then

$$ct=D+vt, \quad ct'=D-vt';$$

whence

$$t=\frac{D}{c-v}, \qquad t'=\frac{D}{c+v}.$$

The whole time required for the passage of the light in both directions is

$$t+t'=\frac{2cD}{c^2-v^2},$$

and the distance traveled in this time is

$$(t+t')c=2D\frac{c^2}{c^2-v^2}=2D\left(1+\frac{v^2}{c^2}\right),$$

the terms of fourth order and higher being neglected in the last member.

The length of the path ABA' is evidently

$$2D\sqrt{1+\frac{v^2}{c^2}}=2D\left(1+\frac{v^2}{2c^2}\right),$$

to the same degree of accuracy as before. The difference of the two lengths is, therefore, approximately Dv^2/c^2.

If the whole apparatus is now turned through an angle of $90°$, the difference will be in the opposite direction, and hence

the displacement of interference fringes along $A'T$ should be $2Dv^2/c^2$.

This is a very small difference even when v is the velocity of the earth in its orbit; but it is altogether sufficient to be detected and measured if it were present with no other effect to mask it. The result of the experiment was that practically no displacement of interference fringes was observed; at most the displacement was less than one-fortieth of that expected.

The conclusions which are to be drawn from this experiment we shall state in the next chapter.

§ 6. Other Experimental Investigations.

On the electrical side the problem of detecting the movement of the earth through the ether has been attacked by Trouton and Noble.* They hung up an electrical condenser by a torsion wire and looked for a torque the presence of which was predicted on the hypothesis of a stationary ether through which the condenser was carried by the motion of the earth. Although the sensitiveness of their electrical arrangement was ample for the observation of the expected effect, no evidence of it was found.

Therefore both on the optical and on the electrical side the attempt to detect the motion of the earth through the ether fails; no experiment is known by which it can be put in evidence.

In addition to this negative evidence concerning the predicted effect of the earth's motion through the ether there is also the positive evidence which comes from the verification of contrary predictions based on other principles. This will come in incidentally for discussion in our later chapters, and consequently will be dismissed here.

The experiments which we have described (and others related to them) are fundamental in the theory of relativity. The postulates of the next chapter are based on them. These postulates are in the nature of generalizations of the facts established by the experiments.

* Phil. Trans. Roy. Soc. (A), 202 (1904): 165.

§ 7. The Theory of Relativity is Independent of the Ether.

In the next chapter we shall begin the systematic development of the theory of relativity. It will be seen that its fundamental postulates, or laws, are based on the experiments of which we have given a brief account and on others related to them. These experiments have been carried out to test predictions which have been made on the basis of a certain theory of the ether. But the results which have been obtained are of a purely experimental character and can be formulated so as not to depend in any way on a theory of the ether. In other words, the laws stated in the postulates in the next chapter are in no way dependent for their truth on either the existence or the non-existence of the ether or on any of its properties.

It is important to keep this in mind on account of the confusion which has sometimes arisen as to the relation between the theory of relativity and the theory of the ether. The postulates, as we shall see, are simply generalizations of experimental facts; and, unless an experiment can be devised to show that these generalizations are not legitimate, it is natural and in accordance with the usual procedure in science to accept them as "laws of nature." They are entirely independent of any theory of the ether.

CHAPTER II.

THE POSTULATES OF RELATIVITY.

§ 8. Introduction.

There are two fundamental postulates concerning the nature of space and time which underlie all physical theory. They assert in part that every point of space is like every other point and that every instant of time is like every other instant. To make the statement of these properties more exact and complete we may say that space is isotropic and homogeneous and three-dimensional, while time is homogeneous and one-dimensional. One important mathematical meaning of this is that the transformations of the space and time coordinates are to be linear.

All our theorems will depend directly or indirectly on these two postulates concerning the nature of space and time. Since it is certain that no one will be disposed seriously to call them in question, it is considered unnecessary to give any further statement of them or to make explicit reference to them as part of the basis on which any particular theorem depends, it being understood once for all that they underlie all our work. ?

In the previous chapter we gave some account of the experiments of Michelson and Morley and of Trouton and Noble. There are different points of view from which one may look at these experiments. In the theory of relativity they are taken in the light of an attempt to detect the earth's motion through space by means of the effect of this motion on terrestrial phenomena. So far as the experiments go, they indicate that such motion cannot be detected in this way. Furthermore, no one has yet been able to devise an experiment by means of

which the earth's motion through space can be detected by observations made on the earth alone.

The question arises: Is it possible to have any such experiment at all? In the theory of relativity this question is answered in the negative. The Michelson-Morley experiment and other experiments have been further generalized into the hypothesis that it is impossible to detect motion through space as such; that is, that the only motion of which we can have any knowledge is the motion of one material body or system of bodies relative to another. A sharp formulation of this conclusion constitutes the first characteristic postulate of relativity.

§ 9. Systems of Reference.

Before stating the postulate, however, it will be necessary to introduce a definition. In order to be able to deal with such quantities as are involved in the measurement of motion, time, velocity, etc., it is necessary to have some system of reference with respect to which measurements can be made. Let us consider any set of things consisting of objects and any kind of physical quantities whatever * each of which is at rest with reference to each of the others. Let us suppose that among these objects are clocks, to be used for measuring time, and rods or rules, to be used for measuring length. Such a set of objects and quantities, at rest relatively to each other, together with their units for measuring time and length, we shall call a system of reference.† Throughout the book we shall denote such a system by S. In case we have to deal at once with two or more systems of reference we shall denote them by S_1, S_2, S_3, . . . or by S, S', . . . Furthermore, *it will be assumed that the units of any two systems S_1 and S_2 are such that the same numerical result will be obtained in measuring with*

* As, for instance, charges, magnets, light-sources, telescopes, etc.

† If any number of these objects or quantities are absent we shall sometimes refer to what remains as a system of reference. Thus the system might consist of a single light-source alone.

the units of S_1 *a quantity* L_1 *and with the units of* S_2 *a quantity* L_2 *when the relation of* L_1 to S_1 *is precisely the same as that of* L_2 *to* S_2.

§ 10. The First Characteristic Postulate.

With this definition before us we are now able to state the first characteristic postulate of relativity:

Postulate M. The unaccelerated motion of a system of reference S cannot be detected by observations made on S alone, the units of measurement being those belonging to S.

The postulate, as stated, is a direct generalization from experiment. None of the actually existing experimental evidence is opposed to it. The conviction that future evidence will continue to corroborate it is so strong that objection has seldom or never been offered to this postulate by either the friends or the foes of relativity. No means at present known will enable the observer to detect motion through space or through any sort of medium which may be supposed to pervade space. Furthermore, in every case where the usual theories have predicted the possibility of detecting such motion and where sufficiently exact observations have been made, it has turned out that no such motion was detected. Moreover, one at least of these contradictions of theory—the Michelson-Morley experiment—has been outstanding for a period of twenty-five years and no satisfactory explanation has been offered unless one is willing to accept the law stated in postulate *M* above. It would appear, therefore, that the experimental evidence for the postulate is to be considered of strong character.

§ 11. The Second Characteristic Postulate.

The so-called second postulate of relativity, in the form in which it has frequently been stated,* involves two entirely distinct parts. To the present writer it appears that no inconsiderable part of the difficulty which has been felt concerning

* See postulate $\bar{R}$ below and the remarks which lead up to it.

this second postulate has been due to a failure to perceive the interdependence of these two parts and of postulate M above. Precisely that part of the second postulate to which most objection has been raised is a logical consequence of M and of the other part, the part last mentioned being a statement of a law which for a long time has been accepted by physicists. Consequently, we shall state separately the two parts of the second postulate and bring out with care the interdependence of these and of postulate M above.

The part which we shall give first states a principle which has long been familiar in the theory of light, namely, that the velocity of light is unaffected by the velocity of the source. Stated in exact language this postulate is as follows:

POSTULATE R'. *The velocity of light in free space, measured on an unaccelerated system of reference S by means of units belonging to S, is independent of the unaccelerated velocity of the source of light.*

The law stated in this postulate is a conclusion which follows readily from the usual undulatory theory of light and will therefore be accepted by any one who holds to that theory. But it should be emphasized that R' does not depend for its truth on any theory of light. It is a matter for direct experimental verification or disproof, and this should be made in such a way as to be independent, as far as possible, of all general theories of light, at least insofar as they are not supported by *direct* experimental evidence. So far as the writer is aware, there is no experimental evidence which is undoubtedly opposed to postulate M, while on the other hand there is direct experimental evidence which is believed by some to be definitely in its favor. Tolman,* in particular has considered this matter in relation to the Doppler effect and to the velocity of light from the two limbs of the sun; and has concluded that experiment bears out the postulate. Stewart,† on the other hand, has examined the same experiments and has found an explanation

* *Physical Review*, 31 (1910): 26–40.
† *Physical Review*, 32 (1911): 418–428.

for them in Thomson's electromagnetic emission theory of light. According to Stewart these experiments are in agreement with our postulate M but are opposed to our postulate R'. All other attempted proof or disproof of the postulate appears to be in the same state; it is capable of two interpretations which are directly opposed to each other with respect to their conclusions as to the validity of R'. Thus at present there is no undoubted experimental evidence for or against postulate R'. If the assumption is to be proved at all, either new experiments must be devised or it must be proved by indirect means by showing that it is a consequence of experiment and accepted laws.

Now any one who accepts postulates M and R' will perforce accept also all the logical consequences which necessarily flow from them. Of these logical consequences we shall now prove one which is of great importance in the theory of relativity:

*THEOREM I. The velocity of light in free space, measured on an unaccelerated system of reference S by means of units belonging to S, is independent of the direction of motion of S(MR').**

Since by R' the velocity of light is independent of that of the light-source we may suppose that the light-source belongs to the system of reference S. Now let the velocity of light, as it is emitted from this source in various directions, be observed and tabulated. On account of the homogeneity and isotropy of space mere direction through space will have no effect on these observed velocities; and therefore if they differ at all, the difference will be due to the velocity of S. Now if there were a difference due to the direction of motion of S this difference would put in evidence the motion of S. But by M it is impossible to detect such motion in this way. Hence the observed velocity must be the same in all directions. In other words, it is independent of the direction of motion of S; and thus the theorem is proved.

It is clear, however, that we cannot take the next step and

* Letters attached to a theorem in this way indicate those of the postulates on which the theorem depends.

prove that this observed velocity of light is independent of the numerical value of the velocity of S. To see this clearly, let us suppose that the numerical value of the velocity of S does effect the observed velocity of light. On account of R' it will have the same effect on the observed velocity of light whatever may be the unaccelerated motion of the light-source. Hence, from all possible observations, the experimenter will have only a single datum from which to determine the effect of one phenomenon on another; namely, a datum in which the two phenomena are connected in a certain definite way. It is obvious then that he cannot determine the effect of one of the phenomena on the other; for he can never observe the one without the other being present also and the connection which exists between them is always the same however he may vary the experiment. And if the observer cannot determine an existing effect it is clear that he cannot prove the absence of any effect whatever.

But, *although the absence of this effect cannot be proved, it is probably impossible to conceive any satisfactory way in which it could be present.* Physical intuition is emphatic in asserting that if the direction of the velocity of S has no effect on the observed velocity of light then the numerical value of the velocity of S has no effect on such observed velocity. But this does not constitute a proof. There is in this, however, nothing to invalidate the naturalness of the *assumption* of such independence of the two velocities; in fact, it would be unscientific to make a different assumption (which would necessarily introduce greater complications) unless we were forced to it by unquestioned experimental fact. Accordingly, we shall make the assumption and shall state it as postulate R'':

POSTULATE R''. The velocity of light in free space, measured on an unaccelerated system of reference S by means of units belonging to S, is independent of the numerical value of the velocity of S.

POSTULATE R. The postulate obtained by combining R' and R'' will, for convenience, often be referred to as postulate R.

Now since unaccelerated velocity is completely determined when the numerical value of the velocity and the direction of

the motion are given the truth of the following theorem is an immediate consequence of theorem I and postulate R'':

THEOREM II. The velocity of light in free space, measured on an unaccelerated system of reference S by means of units belonging to S, is independent of the velocity of S (MR).

The second postulate of relativity has usually been stated in a form different from that given above in R' and R'' or R. In fact, the truth of theorem I has often been taken as part of the *assumption* in this postulate, notwithstanding that I can be derived from M and R'. Now, it is precisely the assumption of I that has given most difficulty to some persons. It is believed that a part of this difficulty will disappear in view of the fact that I is here *demonstrated* by means of M and R'.

For the sake of convenience in future discussion one of the customary formulations of the second postulate is appended here. It must be remembered, however, that it is not a separate constituent part of our present body of doctrine but is already contained in M and R, in part directly and in part as a necessary consequence of these postulates.

POSTULATE $\bar{R}$. The velocity of light in free space, measured on an unaccelerated system of reference S by means of units belonging to S, is independent of the velocity of S and of the unaccelerated velocity of the light-source.

From the very nature of the postulate R'' it is difficult to obtain direct experimental evidence for or against it. It seems, however, as we have previously pointed out, that one who accepts theorem I can hardly refuse to assume R''. But theorem I is a logical consequence of postulates M and R', as we have shown. Moreover, from what follows it will be seen that we have occasion to make no further assumptions which can in any way run counter to currently accepted notions. Consequently, it would seem that the experimental evidence for or against the whole theory of relativity must center around postulates M and R'. We have already given some account of the experimental evidence for these postulates. In connection with theorems to be derived later further reference will

be given to the existing experimental evidence and some other possible lines of research in this direction will be pointed out.

It is generally conceded that the strange conclusions which are obtained in the theory of relativity are due to postulate R (or to postulate $\bar{R}$ in the customary formulation). In view of theorem I above and the discussion of its consequences, it is now clear that the strangeness in the conclusions of relativity is due to that part of R which is contained in R'. It is important therefore to have a careful analysis of this postulate and especially to know alternative forms, which, in view of the other postulates, are logically equivalent to it. We shall return to this matter in Chapter VI.

§ 12. THE POSTULATES V AND L.

It has been customary for writers on relativity to state explicitly only the postulates M and R. But every one, as a matter of fact, has made further assumptions concerning the relations of the two systems. These assumptions in some form are essential to the initial arguments and to the conclusions which are drawn by means of them. To the present writer it seems preferable to have these assumptions explicitly stated. Among several forms, any one of which might be chosen, there is one which seems to be decidedly simpler than any of the others; and it is this one which we shall employ here. We state the postulates V and L as follows:

POSTULATE V. If the velocity of a system of reference S_2 relative to a system of reference S_1 is measured by means of the units belonging to S_1 and if the velocity of S_1 relative to S_2 is measured by means of the units belonging to S_2 the two results will agree in numerical value.

This velocity we shall call the *relative velocity* of the two systems. The direction line of this velocity will be called the *line of relative motion* of the two systems.

POSTULATE L. If two systems of reference S_1 and S_2 move with unaccelerated relative velocity and if a line segment l is per-

pendicular to the line of relative motion of S_1 *and* S_2 *and is fixed to one of these systems, then the length of* l *measured by means of the units belonging to* S_1 *will be the same as its length measured by means of the units belonging to* S_2.

The essential content of these two postulates may be stated in simpler terms (but less accurately) if one allows the explicit introduction of the observer. Thus V is roughly equivalent to the following statement: *Two observers whose relative motion is uniform will agree in their measurement of that uniform relative motion.* As an approximate equivalent of L we have: *Two observers whose relative motion is uniform will agree in their measurement of length in a line perpendicular to their line of relative motion.*

It will be observed that these two postulates are nothing more than explicit statements of notions which underlie the classic theories of mechanics. The first is assumed in supposing that there exists such a thing as the relative motion of two bodies which are not at rest relatively to each other. The second is nothing more than the statement of a portion of the idea which lies at the bottom of our conception of such a thing as the length of a rod or other object.

Since these two postulates are universally accepted, the question might naturally arise, Why state them at all? Is it not enough simply to take them for granted? The answer is that there are other notions which have heretofore met with the same universal acceptance and which do not agree with the postulates of relativity. Therefore it seems to be desirable—in fact, to be essential to proper logical procedure—to state explicitly just those *assumptions* concerning the relation of the two systems of reference which we shall have occasion to employ in argument. Only in this way is one able to see exactly on what basis our strange conclusions rest.

We shall make a digression here to say one further word about postulate L. In the next chapter we shall draw the conclusion that length in the line of motion is not independent of the velocity with which the system is moving. In view of

this the question arises as to why we must assume that length in a line perpendicular to the line of motion is independent of the motion. The answer is that we are under no such necessity, that we are at liberty to assume that length in a line perpendicular to the line of motion is dependent on the velocity of such motion. In fact, the general formulation of such an hypothesis has already been made by E. Riecke.* This hypothesis, however, is undoubtedly more complicated and less elegant than the one which we have made; and the latter, as we shall see, is in conflict with no known experimental facts. Therefore, following that instinct which has always wisely guided the physicist, we make the simplest hypothesis which is in agreement with and explanatory of the totality of experimental facts at present known. If at any time experiments are set forth which do not agree with the theory developed on the basis of the above postulates, then will be the time to consider the question of introducing a more complicated postulate in place of our postulate *L* above.

§ 13. Consistency and Independence of the Postulates.

Throughout our treatment it will be assumed that the postulates as stated above are consistent; that is to say, no attempt will be made to prove their consistency. The fact that no contradictory conclusions have been drawn from them will be accepted as (partial) evidence that they are mutually consistent. Moreover, from their very nature and from the differing range of applicability of the several postulates it is difficult to conceive how any of them can possibly contradict conclusions which may be drawn from the others.

There is another question also which it is our purpose to pass over without discussion, namely, the question of the logical independence of the postulates. Is any postulate or a part of any postulate a logical consequence of the remaining postulates? This question is important from the point of view

* Göttinger Nachrichten, Math. Phys., 1911, pp. 271–277.

of formal logic, but in the present case its value to physical science is probably small.

§ 14. Other Postulates Needed.

From the postulates stated above it is possible to draw only those conclusions of the theory of relativity which are of a general nature and have to do merely with the measurement of time and space. They alone are employed in Chapters III and IV.

If it is desired to study the nature of mass or the relation of mass and energy in the theory of relativity, it is necessary to have some assumption concerning mass in the first case and concerning both mass and energy in the second case. Thus we might assume the conservation laws of energy, electricity and momentum and deduce the joint consequences of these assumptions and those given above. It is our purpose to take up these matters in Chapters V and VI. It is convenient to state the postulates here; and this we do, after giving some necessary definitions.

If m, M and v are respectively the mass, momentum and velocity of a body we shall assume (as in the classical mechanics) that they are connected by a relation of the form

$$M = mv.$$

We shall take mass and velocity to be the fundamental quantities and shall define momentum in terms of them by the above relation.

Likewise we shall define the kinetic energy E of a moving body by means of the usual relation

$$E = \int_0^v M\,dv = \int_0^v mv\,dv.$$

Later we shall see that " mass " is variable and is not in general independent of the direction in which it is measured; consequently, we must take for m in the above formulæ the mass of the body in the direction of its motion.

We shall take for granted the following laws of conservation of momentum and energy and electricity:

POSTULATE C_1. *The sum total of momentum in any isolated system remains unaltered, whatever changes may take place in the system, provided that it is not affected by any forces from without.*

POSTULATE C_2. *The sum total of energy in any isolated system remains unaltered, whatever changes may take place in the system, provided that it is not affected by any forces from without.*

POSTULATE C_3. *The sum total of electricity in any isolated system remains unaltered, whatever changes may take place in the system, provided that the system as a whole neither receives electricity from nor gives out electricity to bodies not belonging to the system.*

The "action" of a moving body in passing from one position to another may be defined as the space integral of the momentum taken over the path of motion. If we denote this action by A we have therefore

$$A=\int Mds=\int mvds.$$

Now $ds=vdt$, so that we have also

$$A=\int mv^2dt.$$

If several bodies are involved we have

$$A=\sum\int mvds=\sum\int mv^2dt,$$

where the summation is for the various bodies in the system.

We may state the fundamental principle of least action in the following form:

PRINCIPLE OF LEAST ACTION. The free motion of a conservative system between any two given configurations has the property that the action A is a minimum, the admissible values $\bar{A}$ of the action with which A is compared being obtained from varied motions in which the total energy has the same constant value as in the actual free motion.

CHAPTER III.

THE MEASUREMENT OF LENGTH AND TIME.

§ 15. Relations between the Time Units of Two Systems.

Let us consider three systems of reference S, S_1 and S_2 related to each other in the following manner: The lines of relative motion of S and S_1, of S and S_2, of S_1 and S_2 are all parallel; S_1 and S_2 have a relative velocity * v; S and S_1 have a relative velocity $\frac{1}{2}v$ in one sense and S and S_2 have a relative velocity $\frac{1}{2}v$ in the opposite sense. The system S consists of a single light-source, and this source is symmetrically placed with respect to two points of which one is fixed to S_1 and the other is fixed to S_2. This is possible as a permanent relation on account of the relative motions of the three systems. For convenience, let us assume S to be at rest.

We shall now suppose that observers on the systems S_1 and S_2 measure the velocity of light as it emanates from the source S. Let a point A on S_1 and a point B on S_2, which are symmetrically placed with respect to the light-source S, move along the lines l_1 and l_2; these lines are parallel. From postulate L it follows that the observers on S_1 and S_2 will obtain the same measurement of the distance between l_1 and l_2. Denote this distance by d. From postulate M it follows that neither observer is able to detect his motion. Therefore he will make his observations on the assumption that his system is at rest; that is to say, his measurements will be made by means of the units

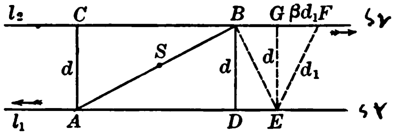

Fig. 3.

* Note that postulate V is required to make this hypothesis legitimate.

belonging to his system and no corrections will be made on account of the motion of the system. Let the observer on S_1 reflect a ray of light SA from a point A to a point C on l_2 and back to A; and let the observed time of passage of the light from A to C and back to A be t. Since the observer assumes his system to be at rest he will suppose that the ray of light passes (in both directions) along the line AC which is perpendicular to l_1 and l_2. His measurement of the distance traversed by the ray of light in time t will therefore be $2d$. Hence he will obtain as a result

$$\frac{2d}{t}=c,$$

where c is his observed velocity of light.

Similarly, an observer on S_2, supposing his system to be at rest finds the time t_1 which it requires for a ray of light to pass from B to D and return, the ray employed being gotten by reflecting a ray SB at B. Thus the second observer obtains

$$\frac{2d}{t_1}=c_1,$$

where c_1 is his observed velocity of light.

Now, from the assumed relations among the systems S, S_1 and S_2 and from the homogeneity of space it follows that the two observations which we have supposed to be made must lead to the same estimate for the velocity of light. This is readily seen from the fact that the observations were made in such a way that the effect due to either the numerical value or the direction of the motion of the systems S_1 and S_2 is the same in the two cases. In other words, if we denote by L_1 and L_2 the quantities measured on S_1 and S_2 respectively, then the relation of L_1 to S_1 is precisely the same as that of L_2 to S_2; and hence the numerical results are equal, as one sees from the definition of systems of reference. Therefore we have $c_1=c$.

Let us now suppose that the observer at A is watching the experiment at B. To him it appears that B is moving with a velocity v, since by hypothesis the two systems have the relative velocity v and A and B measure this velocity alike. We shall

assume that the apparent motion is in the direction indicated by the arrow in the figure. To the observer at B it appears that the ray of light traverses BD from B to D and returns along the same line to B. To the observer at A it appears that the ray traverses the line BEF, F being the point which B has reached by the time that the ray has returned to the observer at this point. If EG is perpendicular to l_2 and d_1 is the length of EF as measured by means of units belonging to S_1, then, evidently, GF (when measured in the same units) is βd_1, where $\beta = v/\bar{c}$ and $\bar{c}$ is the (apparent) velocity of light as estimated in this case by the observer at A. From the right triangle EFG it follows at once that we have

$$d_1 = \frac{d}{\sqrt{1-\beta^2}}.$$

Now, if $\bar{t}$ is the time which is required, according to the observer at A, for the light to traverse the path BEF, then we have

$$\frac{2d_1}{\bar{t}} = \frac{2d}{\bar{t}\sqrt{1-\beta^2}} = \bar{c}.$$

So far in our argument in this section we have employed only those of our postulates which are generally accepted by both the friends and the foes of relativity. Now we come to the place where the men of the two camps must part company.

Let us introduce for the moment the following additional hypothesis:

ASSUMPTION A. *The two estimates c and $\bar{c}$ of the velocity of light obtained as above by the observer at A are equal.*

Now we have shown that c is equal to c_1. Hence we may equate the values of c_1 and $\bar{c}$ given above; thus we have

$$\frac{2d}{t_1} = \frac{2d}{\bar{t}\sqrt{1-\beta^2}};$$

or

$$t_1 = \bar{t}\sqrt{1-\beta^2}.$$

But t_1 and $\bar{t}$ are measures of the same interval of time, t_1 being in units belonging to S_2 and $\bar{t}$ being in units belonging to S_1. Hence to the observer on S_1, the ratio of his time unit to that of

the system S_2 appears to be $\sqrt{1-\beta^2} : 1$. On the other hand, it may be shown in exactly the same way that to the observer on S_2 the ratio of his time unit to that of the system S_1 appears to be $\sqrt{1-\beta^2} : 1$. That is, the time units of the two systems are different and each observer comes to the same conclusion as to the relation which the unit of the other system bears to his own.

This important and striking result may be stated in the following theorem:

THEOREM III. If two systems of reference S_1 and S_2 move with a relative velocity v and β is defined as the ratio of v to the velocity of light estimated in the manner indicated above, then to an observer on S_1 the time unit of S_1 appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of S_2 while to an observer on S_2 the time unit of S_2 appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of S_1 $(MVLA)$.

Let us now bring into play our postulate R'. In theorem I we have already seen that a logical consequence of M and R' is that the velocity of light, as observed on a system of reference, is independent of the direction of motion of that system. Now, if c and $\bar{c}$ as estimated above differ at all, that difference can be due only to the direction of motion of S_1, as one sees readily from postulate R' and the method of determining these quantities. Hence the statement which we made above as assumption A is a logical consequence of postulates M and R'. Therefore we are led to the following corollary of the above theorem:

COROLLARY. Theorem III may be stated as depending on $(MVLR')$ *instead of on* $(MVLA)$.

Let us now go a step further and employ postulate R''. From theorem I and postulates R' and R'' it follows that the observed velocity of light is a pure constant for all admissible methods of observation. If we make use of this fact the preceding result may be stated in the following simpler form:

THEOREM IV. If two systems of reference S_1 and S_2 move with a relative velocity v and β is the ratio of v to the velocity of light, then to an observer on S_1 the time unit of S_1 appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of S_2 while to an observer on S_2 the time

unit of S_2 appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of S_1 (MVLR).

Let us subject these remarkable results to a further analysis. Theorem III, its corollary and theorem IV all agree in the extraordinary conclusion that the time units of the two systems of reference S_1 and S_2 are of different lengths. Just how much they differ is a secondary matter; that they differ at all is the surprising and important thing. As postulates M, V, L are generally accepted and have not elsewhere led to such strange conclusions it is natural to suppose that the strangeness here is not due to them.

Referring to the argument carried out above, we see that no unusual conclusions were reached until we had introduced and made use of assumption A. Moreover we have seen that this assumption itself is a logical consequence of M and R'. Further, R'' is not involved either in theorem III or in its corollary. But these already contain the strange features of our results. Hence the conclusion is irresistible that the extraordinary element in these results is due to postulate R'— or to speak more accurately, to just that part of it which it is necessary to use in connection with M in order to prove A as a theorem.

This result is important, as the following considerations show. Postulates V and L state laws which have been universally accepted in the classical mechanics. Postulate M is a direct generalization from experiment, and the generalization is legitimate according to the usual procedure of physicists in like situations. Postulate R' is a statement of a principle which has long been familiar in the theory of light and has met with wide acceptance. Thus we see that no one of these postulates, in itself, runs counter to currently accepted physical notions. And yet just these postulates alone are sufficient to enable us to conclude that corresponding time units in two systems of reference are of different magnitude. In the next section we shall show on the basis of the same postulates that the corresponding units of length in the two systems are also different. Thus the most remarkable elements in the

conclusions of the theory of relativity are deducible from postulates M, V, L, R' alone; and yet these are either generalizations from experiment or statements of laws which have usually been accepted. Hence we conclude: *The theory of relativity, in its most characteristic elements, is a logical consequence of certain generalizations from experiment together with certain laws which have for a long time been accepted.*

One other remark, of a totally different nature, should be made with reference to the characteristic result of theorem IV. It has to do with the relation between the time units of the two systems. This relation is intimately associated with the fact that each observer makes his measurements on the hypothesis that his own system is at rest, while the other system is moving past him with the velocity v. If both observers should agree to call S fixed and if further in this modified "universe" our postulates V, L, R were still valid it would turn out that the two observers would find their time units in agreement. But, in view of M, the choice of S as fixed would undoubtedly seem perfectly arbitrary to both observers; and the content of the modified postulate R would be essentially different from that of the postulate as we have employed it. Hence, if we accept R as it stands—or, indeed, even a certain part of it, as we have shown above—we must conclude that the time units in the two systems are not in agreement, in fact, that their ratio is that stated in the theorems above.

§ 16. Relations between the Units of Length of Two Systems.

Let us consider three systems of reference S, S_1 and S_2 related in the same manner as in the preceding section except that now the two lines l_1 and l_2 coincide. We suppose that S_1 is moving in the direction indicated by the arrow at A and that S_2 is moving in the direction indicated by the arrow at B.

We suppose that observers at A and B again measure the velocity of light as it emanates from S, this time in the direction of the line of motion. Each will carry out his observations

on the supposition that his system is at rest, for from M it follows that he cannot detect the motion of his system. The observer at A measures the time t_1 of passage of a ray of light from A to C and return to A, the length of AC being d when the measurement is made with a unit belonging to S_1. Likewise, the observer at B measures the time t_2 of passage of a ray of light from B to D and return to B, the length BD being d when measured with a unit belonging to S_2.

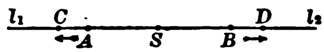

FIG. 4.

Just as in the preceding case it may be shown that the two observers must obtain the same estimate for the velocity of light. But the estimate of the observer at A is $2d/t_1$ while that of the observer at B is $2d/t_2$. Hence

$$t_1 = t_2;$$

that is, the number of units of time required for the passage of the ray at A and of the ray at B is the same, the former being measured on S_1 and the latter on S_2. Moreover, the measure of length is the same in the two cases. But the units of time, as we saw in the preceding section, do not have the same magnitude. Hence the units of length of the two systems along their line of motion do not have the same magnitude; and the ratio of units of length is the same as the ratio of units of time.

Combining this result with theorem III, its corollary and theorem IV we have the following three results:

THEOREM V. If two systems of reference S_1 and S_2 move with a relative velocity v and β is defined as the ratio of v to the velocity of light estimated in the manner indicated in the first part of §15, then to an observer on S_1 the unit of length of S_1 along the line of relative motion appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of S_2 while to an observer on S_2 the unit of length of S_2 along the line of relative motion appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of $S_1(MVLA)$.

COROLLARY. Theorem V may be stated as depending on $(MVLR')$ instead of on $(MVLA)$.

Theorem VI. If two systems of reference S_1 and S_2 move with a relative velocity v and β is the ratio of v to the velocity of light, then to an observer on S_1 the unit of length of S_1 along the line of relative motion appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of S_2 while to an observer on S_2 the unit of length of S_2 along the line of relative motion appears to be in the ratio $\sqrt{1-\beta^2} : 1$ to that of S_1 (MVLR).

We might make an analysis of these results similar to that which we gave for the corresponding results in the preceding section. But it would be largely a repetition. It is sufficient to point out that the remarkable conclusions as to units of length in the two systems rest on just those postulates which led to the strange results as to the units of time.

§ 17. Discussion of the Notion of Length.

In the preceding section we saw that two observers A and B on relatively moving systems of reference S_1 and S_2 respectively are in a very peculiar disagreement as to units of length along a line l parallel to their line of relative motion. To A it appears that B's units are longer than his own. On the other hand, it seems to B that his units are shorter than A's. In the two cases the apparent ratio is the same; more precisely, the unit which appears to either observer to be the shorter seems to him to have the ratio $\sqrt{1-\beta^2} : 1$ to that which appears to him to be the longer. Although they are thus in disagreement there is yet a certain symmetry in the way in which their opinions diverge.

Let us suppose that these two observers now undertake to bring themselves into a closer agreement in measurements of length along the line l. Suppose that B agrees arbitrarily to shorten his unit so that it will appear to A that the units of A and B are of the same length. Then, so far as A is concerned, all difficulty has disappeared. How is B affected by this change? We see that the difficulty which he experienced is not disposed of; on the other hand it is greater than before. Already, it seemed to him that his unit was shorter than A's. Now, since

he has shortened his unit, the divergence appears to him to be increased. Moreover, the symmetry which we found in the former case is now absent.

Furthermore, if any other changes in the units of A and B are made we shall always find difficulties as great as or greater than those which we encountered in the initial case. There is no other conclusion than this: We are face to face with an essential difficulty—one that is not to be removed by any mere artifice. What account of it shall we render to ourselves?

This much is already obvious: The length of an object depends in an essential way upon the measurer and the system to which he belongs.

We have certain intuitive notions concerning the nature of matter which it is necessary for us to examine if we are to discuss adequately the notion of length. We have usually supposed that to revolve a steel bar, for instance, through an angle of ninety degrees has no effect upon its length. Let us suppose for the moment that this is not so; but that the bar is shorter when pointing in some directions than in others, so that its length is the product of two factors one of which is its length in a certain initial position and the other of which is a function of the direction in which the body points relative to that in the initial position. Suppose that at the same time all other objects experience precisely the same change for varying directions. It is obvious that in this case we should have no means of ascertaining this dependence of length upon the direction in which the body points.

To an observer placed in a situation like this it would be natural to assume that the length of the steel bar is the same in all directions. In other words, in arriving at his definition of length he would make certain conventions to suit his convenience.

Now suppose that the system of such an observer is set in motion with a uniform velocity v relative to the previous state of the system; and that at the same time all bodies on his system undergo simultaneously a continuous dilatation or contrac-

tion. This observer would have no means of ascertaining that fact; and accordingly he would suppose that his steel bar had the same length as before. In other words, he would unconsciously introduce a new convention concerning his measurement of length.

There is no *a priori* reason why our actual universe should not be such as the hypothetical one just described. To suppose it so unless our experience demands such a supposition would be unnatural; because it would introduce an unnecessary inconvenience. But suppose that in our growing knowledge of the universe there should come a time when we could more conveniently represent to ourselves the actual facts of experience by supposing that all material things are subject to some such deformations as those which we have indicated above; there is certainly no *a priori* reason why we should not conclude that such is the essential nature of the structure of the universe.

Naturally we would not come to this conclusion without due consideration. We would first enquire carefully if there is not some more convenient way by which we can reconcile all experimental facts; and only in the event of a failure to find such a way would we be willing to modify so profoundly our views of the material world.

Now, if we agree to suppose that our actual universe is subject to a certain (appropriately defined) deformation of the general type discussed above it would follow that observers A and B on the respective systems S_1 and S_2 would be in a disagreement as to units of length similar to that which exists, according to the theory of relativity. Therefore, that which at the outset seemed to be of such essential difficulty is explained easily enough, *if we are willing to modify so profoundly our conception of the nature of material bodies.*

Whether in the present state of science experimental facts demand such a radical procedure is a question which will be answered differently by different minds. To one who accepts the postulates of relativity there is indeed no other recourse; one who refuses to accept them must find some other

satisfactory way to account for experimental facts. The Lorentz theory of electrons gives striking evidence in favor of supposing that matter is subject to some such deformations as those mentioned above; and this evidence is the more important and interesting in that the deformations (as conceived in this theory) were assumed to exist simply in order to be able to account directly for experimental facts.

§ 18. Discussion of the Measurement of Time.*

That two observers in relative motion are in hopeless disagreement as to the measurement of length in their line of relative motion is a conclusion which is probably (at first) sufficiently disconcerting to most of us; but it is an even greater shock to intuition to conclude, as we are forced to do according to the theory of relativity, that there is a like ineradicable disagreement in the measurement of time. A discussion similar to that in the preceding section brings out the fact that our observers A and B cannot possibly arrive at consistent means of measuring intervals of time. The treatment is so far similar to the preceding discussion for length that we need not repeat it; we shall content ourselves with a brief discussion of conclusions to be drawn from the matter.

Why is this inability of A and B to agree in measuring time received in our minds with such a distinct feeling of surprise and shock? It is doubtless because we have such a lively sense of the passage of time. It seems to be a thing which we know directly, and the conclusion in question is contrary to our unsophisticated intuition concerning the nature of time.

But what is it that we know directly? We have an immediate perception of what it is for two conscious phenomena to coexist in our mind, and consequently we perceive imme-

* In connection with this section and the following one the reader should compare the excellent and interesting treatment of the problem of measuring time to be found in Chapter II of Poincaré's Value of Science (translated into English by Halsted).

diately the simultaneity of events in our mind. Further, we have a perfectly clear sense of the order of succession of events in our own consciousness. Is not that all that we know directly?

The difficulties which A and B experience in correlating their measurements of time grow out of two things, of neither of which we have direct perception.

In the first place there are two consciousnesses involved; and what reason have we to suppose that succession of events is the same for these two? This question we shall not treat, assuming that the principal matter can be put into such impersonal form as to obviate this difficulty altogether. (As a matter of fact, so far as anything characteristic of the theory of relativity is concerned this can be done.)

The other difficulty has to do with the measurement of time as opposed to the mere psychological experience of its passage. In this matter we are entirely without any direct intuition to guide us. We have no immediate sense of the equality of two intervals of time. Therefore, whatever definition we employ for such equality will necessarily have in it an important element of convention. To keep this well in mind will facilitate our discussion.

Our problem is this: How shall we assign a numerical measure of length to a given time interval; say to an interval in which a given physical phenomenon takes place? We shall arrive at the answer by asking another question: Why should we seek to measure time intervals at all, seeing that we have no immediate consciousness of the equality of such intervals? There can be only one answer: we seek to measure time as a matter of convenience to us in representing to ourselves our experiences and the phenomena of which we are witnesses. In such a way we can render to ourselves a better account of the world in which we live and of our relation to it.

Now, since our only reason for attempting to measure time is in a matter of convenience, the way in which we measure it will be determined by the dictates of that convenience. The system of time measurement which we shall adopt is just that

system by means of which the laws of nature may be stated in the simplest form for our comprehension.

Let us return to the case of the two observers *A* and *B* of the preceding section. Suppose that each of them has chosen a system of measuring time that suits his convenience in the interpretation of the laws of nature on his system. There is no *a priori* reason why the two observers should measure time intervals in the same way. In fact, since there is an arbitrary element in the case of each method of measurement and since the two systems are in a state of relative motion, it is not at all unnatural that the units of *A* and *B* should differ.

Now it is to be noticed that each of the observers *A* and *B* is in just the situation in which we find ourselves. We have chosen a method of measuring time which seems to us convenient. Insofar as that method depends on convenience it is relative to us who are observers, and therefore it has in it something which is arbitrary. There is no doubt that it would be desirable for us to know what it is which is arbitrary, which is relative to us who observe; but it is equally obvious that it must be difficult for us to determine what this arbitrary element is.

The theory of relativity makes a contribution to the solution of this problem. We suppose that two observers on different systems find the laws of nature the same as we find them; or, more exactly, we suppose that they find certain specific laws the same as we find them. Then we inquire as to their agreement in measuring time and see that they differ in a certain definite way. This difference is due to things which are relative to the two observers; and thus we begin to get some insight into the ultimate basis of our own method of measurement. It is obviously an important service which the theory of relativity renders to us when it enables us to to make an advance towards a better understanding of such a fundamental matter as this.

This matter will become clearer if we speak of the simultaneity of events which happen at different places; and therefore we turn to a discussion of this topic.

§ 19. Simultaneity of Events Happening at Different Places.

Let us now assume two systems of reference S and S' moving with a uniform relative velocity v. Let an observer on S' undertake to adjust two clocks at different places so that they shall simultaneously mark the same hour. We will suppose that he does this in the following very natural manner: Two stations A and B are chosen in the line of relative motion of S and S' and at a distance d apart. The point C midway between these two stations is found by measurement. The observer is himself stationed at C and has assistants at A and B. A single light signal is flashed from C to A and to B, and as soon as the light ray reaches each station the clock there is set at an hour agreed upon beforehand. The observer on S' now concludes that his two clocks, the one at A and the other at B, are simultaneously marking the same hour; for, in his opinion (since he supposes his system to be at rest) the light has taken exactly the same time to travel from C to A as to travel from C to B.

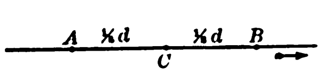

Fig. 5.

Now let us suppose that an observer on the system S has watched the work of regulating these clocks on S'. The distances CA and CB appear to him to be

$$\tfrac{1}{2}d\sqrt{1-\beta^2}$$

instead of $\frac{1}{2}d$. Moreover, since the velocity of light is independent of the velocity of the source, it appears to him that the light ray proceeding from C to A has approached A at the velocity $c+v$, where c is the velocity of light, while the ray going from C to B has approached B at the velocity $c-v$. Thus to him it appears that the light has taken longer to go from C to B than from C to A by the amount

$$\frac{\frac{1}{2}d\sqrt{1-\beta^2}}{c-v}-\frac{\frac{1}{2}d\sqrt{1-\beta^2}}{c+v}=\frac{vd\sqrt{1-\beta^2}}{c^2-v^2}.$$

But since $\beta = v/c$ the last expression is readily found to be equal to

$$\frac{v}{c^2} \cdot \frac{d}{\sqrt{1-\beta^2}}.$$

Therefore, to an observer on S the clocks of S' appear to mark different times; and the difference is that given by the last expression above.

Thus we have the following conclusion:

THEOREM VII. Let two systems of reference S and S′ have a uniform relative velocity v. Let an observer on S′ place two clocks at a distance d apart in the line of relative motion of S and S′ and adjust them so that they appear to him to mark simultaneously the same hour. Then to an observer on S the clock on S′ which is forward in point of motion appears to be behind in point of time by the amount

$$\frac{v}{c^2} \cdot \frac{d}{\sqrt{1-\beta^2}},$$

where c is the velocity of light and $\beta = v/c$ *(MVLR).*

It should be emphasized that the clocks on S' are in agreement in the only sense in which they can be in agreement for an observer on that system who supposes (as he naturally will) that his own system is at rest—notwithstanding the fact that to an observer on the other system there appears to be an irreconcilable disagreement depending for its amount directly on the distance apart of the two clocks.

According to the result of the last theorem the notion of simultaneity of events happening at different places is indefinite in meaning until some convention is adopted as to how simultaneity is to be determined. In other words, *there is no such thing as the absolute simultaneity of events happening at different places.*

How shall we adjust this remarkable conclusion to our ordinary intuitions concerning the nature of time? We shall probably most readily get an answer to this question by inquir-

ing further: What shall we mean by saying that two events which happen at different places are simultaneous?

First of all it should be noticed that we have no direct sense of what such simultaneity should mean. I have a direct perception of the simultaneity of two events in my own consciousness. I consider them simultaneous because they are so interlocked that I cannot separate them without mutilating them. If two things happen which are far removed from each other I do not have a direct perception of both of them in such way that I perceive them as simultaneous. When should I consider such events to be simultaneous?

To answer this question we are forced to the same considerations as those which we met in the preceding section. There can be no absolute criterion by which we shall be able to fix upon any definition as the only appropriate one. We must be guided by the demands of convenience, and by this alone.

In view of these considerations there is nothing unthinkable about the conclusion concerning simultaneity which we have obtained above. An observer A on one system of reference regulates clocks so that they appear to him to be simultaneous. It is apparent that to him the notion of simultaneity appears to be entirely independent of position in space. His clocks, even though they are separated by space, appear to him to be running together, that is, to be together in a sense which is entirely independent of all considerations of space.

But when B from another system of reference observes the clocks of A's system they do not appear to him to be marking simultaneously the same hour; and their lack of agreement is proportional to their distance apart, the factor of proportionality being a function of the relative velocity of the two systems.

Thus instants of time at different places which appear to A to be simultaneous in a sense which is entirely independent of all considerations of space appear to B in a very different light; namely, as if they were different instants of time, the one preceding the other by an amount directly proportional

to the distance between the points in space at which events occur to mark these instants. Even the order of succession of events is in certain cases different for the two observers, as one can readily verify.

It thus appears that the notion of simultaneity at different places is relative to the system on which it is determined. The only meaning which it can have is that which is given to it by convention.

CHAPTER IV.

EQUATIONS OF TRANSFORMATION.

§ 20. Transformation of Space and Time Coordinates.

It is now an easy matter to derive the Einstein formulæ for the transformation of space and time coordinates. Let two systems of reference S and S' have the relative velocity v in the line l. Let systems of rectangular coordinates be attached to the systems of reference S and S' in such a way that the x-axis of each system is in the line l, and let the y-axis and the z-axis of one system be parallel to the y-axis and the z-axis respectively of the other system. Let the origins of the two systems coincide at the time $t = 0$. Furthermore, for the sake of distinction, denote the coordinates on S by x, y, z, t and those on S' by x', y', z', t'. We require to find the value of the latter coordinates in terms of the former.

From postulate L it follows at once that $y' = y$ and $z' = z$. Let an observer on S consider a point which at time $t = 0$ appears to him to be at distance * x from the $y'z'$-plane; at time $t = t$ it will appear to him to be at the distance $x - vt$ from the $y'z'$-plane. Now, by an observer on S' this distance is denoted by x'. Then from theorem VI we have

$$x'\sqrt{1-\beta^2} = x - vt.$$

Now consider a point at the distance x from the yz-plane at time $t = t$ in units of system S. From theorem VII it follows

* The algebraic sign of the distance is supposed to be taken into account in the value of x.

that to an observer on S the clock on S' at the same distance x from the yz-plane will appear behind by the amount

$$\frac{v}{c^2}x,$$

where c is the velocity of light. That is to say, in units of S this clock would register the time

$$t-\frac{v}{c^2}x.$$

Hence, by means of theorem IV, we have at once the result

$$t'\sqrt{1-\beta^2}=t-\frac{v}{c^2}x.$$

Solving the two equations involving x' and t' and collecting results, we have

$$(A) \qquad \begin{aligned} t' &= \frac{1}{\sqrt{1-\beta^2}}\left(t-\frac{v}{c^2}x\right), \\ x' &= \frac{1}{\sqrt{1-\beta^2}}(x-vt), \\ y' &= y, \\ z' &= z, \end{aligned} \qquad (MVLR)$$

where $\beta=v/c$ and c is the velocity of light.

In the same way we may obtain the equations which express t, x, y, z in terms of t', x', y', z'. But these can be found more easily by solving equations (A) for t, x, y, z. Thus we have

$$(A_1) \qquad \begin{aligned} t &= \frac{1}{\sqrt{1-\beta^2}}\left(t'+\frac{v}{c^2}x'\right), \\ x &= \frac{1}{\sqrt{1-\beta^2}}(x'+vt'), \\ y &= y', \\ z &= z'. \end{aligned} \qquad (MVLR)$$

These two sets of equations (A) and (A_1) are identical in form except for the sign of v. This symmetry in the transformations constitutes one of their chief points of interest.

§ 21. The Addition of Velocities.

We shall now derive the formulæ for the addition of velocities.

Let the velocity of a point in motion be represented in units belonging to S' and to S by means of the equations

$$x'=u_x't', \quad y'=u_y't', \quad z'=u_z't';$$
$$x=u_xt, \quad y=u_yt, \quad z=u_zt,$$

respectively. In the first of these substitute for t', x', y', z' their values given by (A), solve for x/t, y/t, z/t and replace these quantities by their equals u_x, u_y, u_z respectively. Thus we have

$$(B) \qquad \begin{aligned} u_x &= \frac{u_x'+v}{1+\frac{vu_x'}{c^2}}, \\ u_y &= \frac{\sqrt{1-\beta^2}}{1+\frac{vu_x'}{c^2}}u_y', \\ u_z &= \frac{\sqrt{1-\beta^2}}{1+\frac{vu_x'}{c^2}}u_z'. \end{aligned} \qquad (MVLR)$$

From these results it follows that the law of the parallelogram of velocities is only approximate. This conclusion of the theory of relativity has given rise, in the minds of some persons, to the most serious objections to the entire theory.

Suppose that both the velocities considered above are in the line of relative motion of S and S'. Then we have

$$u=\frac{v+u'}{1+\frac{vu'}{c^2}}.$$

This equation gives rise to the following theorem:

Theorem VIII. If two velocities, each of which is less than c, are combined the resultant velocity is also less than c (MVLR).

To prove this we substitute in the preceding equation for v and u' the values

$$v=c-k, \quad u'=c-l$$

where each of the numbers k and l is positive and less than c. Then the equation becomes

$$u=c\frac{2c-k-l}{2c-k-l+\dfrac{kl}{c}}.$$

The second member is evidently less than c. Hence the theorem.

If, however, either one (or both) of the velocities v and u' is equal to c—and hence k or l (or both) is equal to zero—we see at once from the last equation that $u=c$. Hence, we have the following result:

THEOREM IX. *If a velocity c is compounded with a velocity equal to or less than c, the resultant velocity is c(MVLR).*

§ 22. Maximum Velocity of a Material System.

A conclusion of importance is implicity involved in the preceding results. It can probably be seen in the simplest way by reference to the first two equations (A), these being nothing more nor less than an analytic formulation of theorems IV and VI. If β is in numerical value greater than 1—whence $1-\beta^2$ is negative—the transformation of time coordinates from one system to the other gives an imaginary result for the time in one system if the time in the other system is real. Likewise, measurement of length in the direction of motion is imaginary in one system if it is real in the other. Both of these conclusions are absurd and hence the numerical value of β is equal to or less than 1. If it is 1, then any length in one system, however short, would be measured in the other as infinite; and a like result holds for time. Hence β is numerically less than 1. But $\beta=v/c$, the ratio of the relative velocity of the two systems to the velocity of light. Hence:

THEOREM X. The velocity of light is a maximum which the velocity of a material system may approach but can never reach (MVLR).

It should be pointed out that this theorem may also be proved directly by means of theorem IX.

§ 23. TIME AS A FOURTH DIMENSION.

I have no intention of asserting that time is a fourth dimension of space in the sense in which we ordinarily employ the word "dimension"; such a statement would have no meaning. I wish to point out rather that it is in some measure connected with space, and that in many formulæ it must enter as it would if it were essentially and only a fourth dimension.

We shall see this readily if we examine the formulæ (A) of transformation from one system of reference to another. Here the time variable t enters in a way precisely analogous to that in which the space variables x, y, z enter.

Suppose now that the law of some phenomenon as observed on S' is given by the equation

$$F(x', y', z', t') = 0$$

and we desire to know the expression of this law on S. We substitute for x', y', z', t' their values in terms of x, y, z, t given in (A); and thus we obtain an equation stating the law in question.

From these considerations it appears that in many of our problems, namely in those which have to do at once with two or more systems of reference, the time and space variables taken together play the role of four variables each having to do with one dimension of a four-dimensional continuum.

This conclusion raises philosophical questions of profound importance concerning the nature of space and time; but into these we cannot enter here.

CHAPTER V.

MASS AND ENERGY.

§ 24. Dependence of Mass on Velocity.

Suppose that we have two systems of reference S_1 and S_2 moving with a relative velocity v. We inquire as to whether, and in what way, the mass of a body as measured on the two systems depends on v. Will a given body have the same measure of mass when that mass is estimated in units of S_1 and in units of S_2? And will the mass of a body depend on the direction of its motion by means of which that mass is measured? Our purpose in this section is to answer these two questions.

The two most important directions in which to measure the mass of a body are, first, that perpendicular to the line of relative motion of S_1 and S_2, and, secondly, that parallel to this line of motion. For convenience in distinguishing these we shall speak of the " transverse mass " of a body as that with which we have to deal when we are concerned with the motion of the body in a direction perpendicular to the line of relative motion of S_1 and S_2; when the motion is parallel to this line we shall speak of the " longitudinal mass " of the body.

Lewis and Tolman (Phil. Mag. 18: 510–523) determine what they call the " mass of a body in motion," employing for this purpose a very simple and elegant method. This " mass " is what we have just defined as the transverse mass of the body. We employ the excellent method of these authors in deriving the formula for transverse mass.

Suppose that an experimenter A on the system S_1 constructs a ball B_1 of some rigid elastic material, with unit volume, and

puts it in motion with unit velocity in a direction perpendicular to the line of relative motion of S_1 and S_2, the units of measurement employed being those belonging to S_1. Likewise suppose that an experimenter C on S_2 constructs a ball B_2 of the same material, also of unit volume, and puts it in motion with unit velocity in a direction perpendicular to the line of relative motion of S_1 and S_2; we suppose that the measurements made by C are with units belonging to S_2. Assume that the experiment has been so planned that the balls will collide and rebound over their original paths, the path of each ball being thought of as relative to the system to which it belongs.

Now the relation of the ball B_2 to the system S_1 is the same as that of the ball B_1 to the system S_2, on account of the perfect symmetry which exists between the two systems of reference in accordance with previous results. Therefore the change of velocity of B_2 relative to its starting point on S_2 as measured by A is equal to the change of velocity of B_1 relative to its starting point on S_1 as measured by C. Now velocity is equal to the ratio of distance to time: and in the direction perpendicular to the line of relative motion of the two systems the units of length are equal; but the units of time are unequal. Hence to either of the observers the change of velocity of the two balls, each with respect to its starting point on its own system, will appear to be unequal.

To A the time unit on S_2 appears to be longer than his own in the ratio $1 : \sqrt{1-\beta^2}$ (see theorem IV). Hence to A it must appear that the change in velocity of B_2 relative to its starting point is smaller than that of B_1 relative to its starting point in the ratio $\sqrt{1-\beta^2} : 1$. But the change in velocity of each ball multiplied by its mass gives its change in momentum. From postulate C_1 it follows that these two changes of momentum are equal. Hence to A it appears that the mass of the ball B_1 is smaller than that of the ball B_2 in the ratio $\sqrt{1-\beta^2} : 1$.

Similarly, it may be shown that to C it appears that the mass of the ball B_2 is smaller than that of B_1 in the ratio $\sqrt{1-\beta^2} : 1$.

From our general results concerning the measurement of length it follows that if the ball which has been constructed by A were transferred to C's system it would be impossible for C to distinguish A's ball from his own by any considerations of shape and size. Likewise, as A looks at them from his own system he is similarly unable to distinguish them. It is therefore natural to take the mass of C's ball as that which A's would have if it had the velocity v with respect to S_1 of the system S_2. Thus we obtain a relation existing between the mass of a body in motion and at rest.

Now, "mass" as we have measured it above is the transverse mass of our definition. From the argument just carried out we are forced to conclude that the transverse mass of a body in motion depends (in a certain definite way) on the velocity of that motion. The result may be formulated as follows:

THEOREM XI. *Let m_0 denote the mass of a body when at rest relative to a system of reference S. When it is moving with a velocity v relative to S denote by $t(m_v)$ its transverse mass, that is, its mass in a direction perpendicular to its line of motion. Then we have*

$$t(m_v) = \frac{m_0}{\sqrt{1-\beta^2}},$$

where $\beta = v/c$ and c is the velocity of light ($MVLRC_1$).

In the statement of this theorem we have tacitly assumed that the mass of a body at rest relative to S, when measured by means of units belonging to S, is independent of the direction in which it is measured. If this assumption were not true we should have a means of detecting the motion of S, a conclusion which is in contradiction to postulate M.

In order to find the longitudinal mass of a moving body we first find the relation which exists between longitudinal mass and transverse mass. We employ for this purpose the elegant method of Bumstead (Am. Journ. Science (4) 26: 498–500).

Let us as usual consider two systems of reference S_1 and S_2 moving with a relative velocity v, observers A and B being

stationed on S_1 and S_2 respectively. Suppose that B performs the following experiment: He takes a rod of two units length, whose mass is so small as to be negligible, and attaches to its ends two balls of equal mass. Then he suspends this rod by a wire so as to form a torsion pendulum. We assume that the line of relative motion of the two systems is perpendicular to the line of this wire.

Let us consider the period of this torsion pendulum in the two cases when the rod is clamped to the wire so as to be in equilibrium in each of the following two positions: (1) With its length perpendicular to the line of relative motion of S_1 and S_2; (2) with its length parallel to this line of motion.

As B observes it the period must be the same in the two cases; for, otherwise, he would have a means of detecting his motion by observations made on his system alone, contrary to postulate M. Then from the relation of time units on S_1 and S_2 it follows that the two periods will also appear the same to A. As observed by B the apparent mass of the balls is the same in both cases. We inquire as to how they appear to A. Let m_1 and m_2 be the apparent masses, as observed by A, in the first and second cases respectively. It is obvious that m_1 is the longitudinal mass and m_2 the transverse mass of the balls in question.

When the pendulum is in motion it appears to B that each ball traces a circular arc. From the relations between the units of length in the two systems it follows that to A it appears that the balls trace arcs of an ellipse whose semiaxes are 1 and $\sqrt{1-\beta^2}$ and lie perpendicular and parallel, respectively, to the line of relative motion of the two systems.

Let us now determine the period of each of these two pendulums as they are observed by A. By equating the expressions for these periods we shall find the relation which exists between m_1 and m_2.

Let x and y be the cartesian coordinates of a point as determined by A, the axes of reference being the major and minor axes of the ellipse in which the balls move. Let x' and y' be

the coordinates of the same point as determined by B. Then the circular path of motion, as determined by B, has the equations

$$x'=\cos\theta,\ y'=\sin\theta,$$

the angle θ being measured from the major axis of the ellipse. The equations of the ellipse, as determined by A, are

$$x=\cos\theta,\ y=\sqrt{1-\beta^2}\sin\theta.$$

In the first case—when the rod is perpendicular to the line of relative motion of S_1 and S_2—the amount of twisting in the wire when the ball is in a given position is the numerical value of the corresponding angle θ; and therefore the potential energy * is proportional to θ^2, say that it is $\frac{1}{2}k\theta^2$. Now from the values of y and x above we have

$$y=x\sqrt{1-\beta^2}\tan\theta.$$

For small oscillations we have $x=1$ and $\tan\theta=\theta$; and therefore

$$y=\sqrt{1-\beta^2}\cdot\theta.$$

Hence the potential energy is

$$\tfrac{1}{2}\frac{k}{1-\beta^2}y^2;$$

and the equation of motion of the particle becomes

$$m_1\frac{d^2y}{dt^2}=-\frac{k}{1-\beta^2}y.$$

Hence the period T_1 of oscillation is

$$T_1=2\pi\sqrt{\frac{m_1(1-\beta^2)}{k}}.$$

* That the potential energy is proportional to θ^2 when measured by B is obvious. Since A observes a different apparent angle θ' (say) corresponding to B's observed angle θ it might at first sight appear that the potential energy as observed by A is proportional to θ'^2; that this is not the case is seen from the fact that for a given twist in the wire θ' depends on the direction of equilibrium of the bar, that is, it depends on the way in which the bar is attached to the wire; hence, if the potential energy as observed by A were proportional to θ'^2, it would depend on the way in which the bar is attached. Since this is obviously not the case we conclude that the potential energy is proportional to θ^2.

In the second case—when the rod is parallel to the line of relative motion of S_1 and S_2—the amount of twisting in the wire for a given position of the balls is the numerical value of $\frac{\pi}{2}-\theta$. The potential energy is $\frac{1}{2}k\left(\frac{\pi}{2}-\theta\right)^2$. We have

$$x=\frac{y}{\sqrt{1-\beta^2}}\cot\theta.$$

For small oscillations we have

$$y=\sqrt{1-\beta^2},\quad \cot\theta=\tan\left(\frac{\pi}{2}-\theta\right)=\frac{\pi}{2}-\theta.$$

Hence the potential energy is $\frac{1}{2}kx^2$, and the period T_2 of oscillation is therefore

$$T_2=2\pi\sqrt{\frac{m_2}{k}}.$$

Equating the two periods of oscillation found above we have

$$m_2=(1-\beta^2)m_1.$$

Remembering that m_1 and m_2 are the longitudinal mass and the transverse mass, respectively, and making use of theorem XI, we have the following result:

Theorem XII. Let m_0 denote the mass of a body when at rest relative to a system of reference S. When it is moving with a velocity v relative to S denote by $l(m_v)$ its longitudinal mass, that is, its mass in a direction parallel to its line of motion. Then we have

$$l(m_v)=\frac{m_0}{(1-\beta^2)^{\frac{3}{2}}},$$

where $\beta=v/c$ and c is the velocity of light ($MVLRC_1C_2$).

§ 25. On the Dimensions of Units.

Denote the fundamental measurable physical entities mass, length and time by M, L and T respectively. Then the definition of derived entities gives rise to the so-called dimensional equa-

tions. Thus if V denote velocity we have the dimensional equation

$$V=\frac{L}{T}.$$

That such equations must be useful in obtaining the relations of units in two systems of reference is obvious. Thus from the above dimensional equation for V we may at once derive the fundamental result (see theorem VI) concerning the relation of units of length in the line of relative motion of two systems not at rest relatively to each other. For this purpose it is sufficient to employ postulate V and theorem IV. The reader can easily supply the argument. Or, conversely, if one knows the relations which exist between units of length and units of time in two systems one concludes readily to the truth of postulate V.

Likewise, from the dimensional equation

$$\text{acceleration}=\frac{L}{T^2},$$

one may readily determine the relations which exist between units of acceleration on two systems, it being assumed that the relations of time units and length units are known. Making this assumption, then, the two dimensional equations above give us the following theorem:

Theorem XIII. Let two systems S_1 and S_2 move with a relative velocity v in the direction of a line l, and let $\beta=v/c$ where c is the velocity of light. Then to an observer on S_1 it appears that the unit of velocity [acceleration] on S_1 bears to the unit of velocity [acceleration] on S_2 the ratio $1:1[1:\sqrt{1-\beta^2}]$ or $1:\sqrt{1-\beta^2}$ $[1:1-\beta^2]$ according as the motion is parallel to l or perpendicular to $l(MVLR)$.

Let us use F to denote force. Then from the dimensional equation

$$F=\frac{ML}{T^2},$$

we shall be able to draw an interesting conclusion concerning the measurement of force.

Suppose that an observer B on a system S_2 carries out some observations concerning a certain rectilinear motion, measuring the quantities M', L', T', so that he has the equation

$$F' = \frac{M'L'}{T'^2}.$$

Another observer A on a system S_1 (having with respect to S_2 the velocity v in the line l) measures the same force calling it F. Required the value of F in terms of F', when the motion is parallel to l and when it is perpendicular to l, the estimate being made by A.

When the motion is perpendicular to l—that is, when the force acts in a line perpendicular to l—we have

$$F_1 = \frac{ML}{T^2} = \frac{M'\sqrt{1-\beta^2}\cdot L'}{T'^2(1-\beta^2)} = \frac{F'}{\sqrt{1-\beta^2}}.$$

When the motion is parallel to l we have

$$F_2 = \frac{ML}{T^2} = \frac{M'(1-\beta^2)^{\frac{3}{2}}\cdot L'\sqrt{1-\beta^2}}{T'^2(1-\beta^2)} = (1-\beta^2)F'.$$

These results may be stated in the following theorem:

Theorem XIV. In the same systems of reference as in theorem XIII, let an observer on S_2, measure a given force F' in a direction perpendicular to l and in a direction parallel to l, and let F_1 and F_2 be the values of this force as measured in the first and second cases respectively by an observer on S_1. Then we have

$$F_1 = \frac{F'}{\sqrt{1-\beta^2}}, \quad F_2 = (1-\beta^2)F' \quad (MVLRC_1C_2).$$

It is obvious that a similar use may be made of the dimensional equation of any derived unit in determining the relation which exists between this unit in two relatively moving systems of reference.

§ 26. Mass and Energy.

If, as is frequently done, we employ for the definition of the kinetic energy E the relation (compare § 14)

$$E=\int_0^v \dot{M}\, dv=\int_0^v mv\,dv,$$

it is clear that for the mass m we should take the longitudinal mass $l(m_v)$. Then let m_0 denote the mass of the body at rest, E_0 its energy when at rest (that is, the energy due to its internal activity), and E_v its energy when moving at the velocity v. Then clearly $E=E_v-E_0$, so that in view of theorem XII we have

$$E=E_v-E_0=\int_0^v \frac{m_0 v\,dv}{(1-\beta^2)^{\frac{3}{2}}};$$

whence, on integration, we have

$$E=E_v-E_0=m_0c^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right). \qquad (1)$$

Hence for the kinetic energy of a moving body we have

$$E=m_0c^2(\tfrac{1}{2}\beta^2+\tfrac{3}{8}\beta^4+\ \ldots\ );$$

or, to a first approximation only,

$$E=\tfrac{1}{2}m_0v^2.$$

Therefore *the usual formula for kinetic energy in the classical mechanics is only a first approximation.*

Since relation (1) is to be true for all values of v it is obvious that we have

$$E_v=\frac{m_0c^2}{\sqrt{1-\beta^2}}+k, \quad E_0=m_0c^2+k,$$

where k is a constant, that is, a quantity independent of v. From the first of these equations we conclude further that

$$E_v=c^2\cdot t(m_v)+k$$

so that the total energy of a body, decreased by the constant k, is directly proportional to its transverse mass. In case the body is at rest its mass in one direction is the same as in another;

hence $m_0 = t(m_0)$. Bearing this in mind we have the following theorem:

THEOREM XV. Let m_0 be the mass of a body when at rest with respect to a given system of reference and let $t(m_v)$ denote its transverse mass when it is moving with the velocity v (the case $v=0$ is not excluded). Then the total energy E_v which it possesses is $c^2 \cdot t(m_v) + k$, where k is a constant.

§ 27. ON MEASURING THE VELOCITY OF LIGHT.

The following relations are immediate consequences of equations written out above:

$$\frac{E_v - E_0}{t(m_v) - m_0} = c^2, \quad E_v - k = \frac{E_0 - k}{\sqrt{1 - \beta^2}}.$$

Now, suppose that an experimenter contributes to a body which is at rest a known amount of energy and determines the velocity which this causes the body to acquire. If the two measurements are made with sufficient accuracy one will be able, by substituting the results in the first of the above equations, to determine in this way the velocity of light. Actually to carry out this remarkable method for measuring c would doubtless be very difficult; but the obvious great importance of the result is certainly such as to justify a careful consideration of the problem. If the value of c determined in this way should agree well with its value as otherwise found, this would give us an interesting confirmation of the theory of relativity.

Let us consider the mass of a rotating top, the mass being measured along the axis of rotation. According to our results this mass should be different from that of the same top when at rest, and the difference should bear a definite relation to the amount of energy which is involved in the rotation. If the measurements here involved could be made with sufficient accuracy we would have another means, independent of light itself, for the measurement of the light-velocity c. Again, this experiment would afford us a measure of transverse mass

and in that way could lead to a confirmation of the theory of relativity, provided that we assume c as known from independent measurements; and this confirmation, it is to be noticed, would be independent of electrical considerations.

Remark.—It seems to be impossible to determine the constant k which enters into the above discussion. But in the absence of any evidence to the contrary it would appear natural tentatively to assume that k is zero. On the basis of this assumption we should have the following remarkable conclusions: The mass of a body at rest is simply the measure of its internal energy. The transverse mass of a body in motion is the measure of its internal energy and its kinetic energy taken together. Its longitudinal mass is its total energy multiplied by a simple factor. One can hardly resist the conclusion that the transverse mass of a body depends entirely on its energy, and therefore that matter is merely one manifestation of energy.

§ 28. On the Principle of Least Action.

In § 26 we saw that in the theory of relativity the classical formula $E=\frac{1}{2}mv^2$ for the measure of kinetic energy is true only as a first approximation. This is due to the fact that mass is a variable quantity. But the conclusion does not appear to necessitate our surrender of the law of conservation of energy.

The same causes which lead to a modification in the formula for E will also require a corresponding modification in the value of the action A as defined in § 14. The question arises as to whether the principle of least action is left intact. I cannot enter upon the investigation here; but the problem seems to me to be of importance and consequently I am stating it in the hope that some one will be led to consider the solution.

Undoubtedly the principle of least action is one which should be given up only when there are strong reasons for it. It is a mathematical formulation of the law that nature accomplishes her ends with the least expenditure of labor, so to

speak. Certainly this law is one which appeals to our minds with strong force. There is something about it which is aesthetically satisfying in a high degree. It seems to me, however, that a fresh study of it should be made in the light of the theory of relativity.

§ 29. A Maximum Velocity for Material Bodies.

There are several ways by which it may be shown that a material body cannot have a velocity as great as that of light. One of these we used in § 22, showing that, if a material body had a velocity greater than that of light, the numerical measure of length and time on that body would be imaginary, while if its velocity were just equal to that of light a given time interval would have an infinite measure.

We may also prove the same theorem by means of a consideration of mass. Let us consider the equation

$$l(m_v) = \frac{m_0}{(1-\beta^2)^{\frac{3}{2}}},$$

where m_0 is the mass of a body at rest relative to a given system of reference S and $l(m_v)$ is the longitudinal mass of the body moving with a velocity v with respect to S. If we consider larger and larger values of the velocity v we see that $l(m_v)$ increases and becomes infinite as v approaches c. This is equivalent to saying that the longitudinal mass of any material body becomes infinite as the velocity of that body approaches c. Therefore it would require an infinite force to give to a material body the velocity c; that is, c is a maximum velocity which the velocity of a material body may approach but can never reach.

§ 30. On the Nature of Mass.

This conclusion concerning the maximum velocity of a material body brings up important considerations concerning the essential nature of mass and material things. How shall we conceive of matter so that it should have this astonishing property?

In the present state of science any answer to this question must necessarily be of a speculative character; but it is probably worth while to mention briefly a theory of mass which is consistent with the existence of a maximum velocity for a material body.

Let us suppose that the mass of a piece of matter is due to a kind of strain in the ether, and that this strain is principally localized in a relatively small portion of space, but that from this center of localization there go out to infinity in all directions lines of strain which belong essentially to the piece of matter. (We make no assumption as to how this strain is set up; it may be due largely or entirely to the motion of electrons in the molecules of the matter.) Suppose that these lines of strain, except in the immediate neighborhood of the center of localization, are of such nature as to escape detection by our usual methods. Suppose further that when the piece of matter is moved, that is, when the center of localization is displaced, these lines of strain have a corresponding displacement, but that the ether of space resists this displacement, the degree of resistance depending on the velocity.

If the mass of matter is due to such a strain in the ether it is natural to suppose that mass is a measure of the amount of that strain. But, on our present hypothesis, we see that when matter is moved through space there is an increase of the strain on the ether due to such motion. This manifests itself to us in the way of an increase in the mass of the given piece of matter.

Moreover, when the body is in motion it is natural to suppose that these lines of strain are not distributed evenly in all directions. On account of this fact it would not be a matter for surprise if the mass of a moving body were different in different directions.

It thus appears that appropriate hypotheses (which have nothing in them inherently unnatural) would lead us to expect the same descriptive properties of mass as those which are actually found to exist if one accepts the postulates of relativity. Hence we conclude that there is nothing *a priori*

improbable in the conclusions of relativity concerning the nature of mass. Therefore if we find satisfactory grounds for accepting the initial postulates of relativity, we shall not throw these postulates overboard because of the strange conclusions concerning mass to which they have led us.

Now, if mass is merely a manifestation of energy in the form of a strain in the ether it would follow that gravitation is simply an interaction among these several strains. A strain principally localized in one place would have lines of strain going out from it in all directions, and the action of these lines of strain upon one another would afford the effective means by which gravitation acts.

§ 31. The Mass of Light.

From some results in the preceding discussion it has appeared that the transverse mass of a body is merely a manifestation of its total energy, that it is in fact a measure of that energy. It is then natural to suppose, on the other hand, that anything which possesses energy has mass; and we thus conceive of mass and energy as coextensive.

Now a beam of light possesses energy; whence we conclude naturally that it also has mass. But we have seen that no " material body " can have a velocity as great as that of light. How are these two facts to be reconciled? If we define " matter " as that which possesses mass (and this is probably the best definition) we shall, as we have seen, perhaps best be able to represent to ourselves the nature of matter if we think of it as a strain in the ether. Then the two facts which we have to reconcile would be entirely consistent if we suppose that the beam of light sets up a strain in the ether (whence its mass) but that this strain as a whole is not propagated with the velocity of light. In fact, if it moves at all it is probably with a velocity much smaller than that of light.

CHAPTER VI.

EXPERIMENTAL VERIFICATION OF THE THEORY.

§ 32. Two Methods of Verification.

There are at least two ways in which it may be possible to demonstrate experimentally the accuracy of the theory of relativity.

The first method is direct. It consists in the proof by experiment of the postulates on which the theory is based. These proved, the whole theory then follows by logical processes alone. In Chapter II we have given a sufficient discussion of this method.

The second method is indirect; it may be described as follows: Among the consequences of the theory of relativity seek out one which has the property that if it is *assumed* the postulates of relativity may themselves then be deduced by logical processes alone. If then this assumption is proved experimentally this is sufficient to establish the postulates of relativity, and hence the whole theory. Or, one may find such experimental results as lead to all the essential conclusions of relativity, whence one naturally concludes to the accuracy of the whole theory. A discussion of proofs of this kind will be given in this chapter.

This indirect method of proof is in many cases open to an objection of a kind which does not obtain in the case of the direct method previously mentioned. In the indirect method some auxiliary law, as for instance the law of conservation of electricity, must usually be employed in deducing the relativity postulates or essential conclusions from the new assumption which one seeks to justify by experiment. There is always

the possibility that the auxiliary law itself is wrong; and consequently one's confidence in the accuracy of the relativity postulates as thus deduced can be no stronger than that in the truth of the auxiliary law. The same objection can also be raised against many conclusions which we are accustomed to accept with confidence.

To many persons it appears that the first method of proof mentioned above has been carried out successfully and satisfactorily. But if one does not share this opinion it is still legitimate to accept the theory of relativity as a working hypothesis, to be proved or disproved by future experiment. It is an historical fact, patent to every student of scientific progress, that many of our fundamental laws have been accepted in just this way. Take, for instance, the law of conservation of energy. There is no experimental demonstration of this law; and in the very nature of things it is hard to see how there could be. On the other hand it is at variance with no known experimental fact. Moreover, it furnishes us a very valuable means of systematizing our known facts and representing them to our minds as an ordered whole. In other words, it is the most *convenient* hypothesis to make in the face of the phenomena which we have observed. Similarly, even if one does not believe that the theory of relativity has been conclusively demonstrated, should he not accept that theory (tentatively at least) provided it furnishes him with the most convenient means of representing external phenomena to his mind?

It should further be said that every supposed proof of the theory of relativity is of such character that objections can be raised to it; likewise every supposed disproof of the theory is in the same state. In the meantime, though we cannot accept the theory with all confidence, we can at least use its conclusions to suggest experiments which otherwise would not have been conceived. Therefore, whether true or false, the theory will be useful in the advancement of science.

§ 33. Logical Equivalents of the Postulates.

In every body of doctrine which consists of a finite number of postulates and their logical consequences there are necessarily certain theorems which have the following fundamental relation to the whole body of doctrine: By means of one of these theorems and all the postulates but one that remaining postulate may be demonstrated. That is, one may *assume* such a theorem in place of one of the postulates and then demonstrate that postulate. When the postulate has thus been proved it may be used in argument as well as the theorem itself; hence it is clear that all of the consequences which were obtained from the first set of postulates may now be deduced again, though perhaps in a somewhat different manner. That is, if we consider the whole body of doctrine, composed of postulates and theorems, this totality is the same in the two cases. Two sets of postulates which thus give rise to the same body of doctrine (consisting of postulates and theorems together) are said to be logically equivalent.

The problem of the logical equivalents of a given set of postulates is readily seen to be an important one. The principal value of such a matter, from the point of view of physical science, consists in the fact that it affords alternative methods for the experimental proof or disproof of a theory and that it emphasizes in an effective way the essential difficulties and limitations of such experimental verification in general.

When the indirect method of demonstration described in § 32 is carried out by means of logical equivalents of the postulates it is not open to the objection mentioned above. Unfortunately, it seems to be difficult to carry it out in this way, and consequently we are forced to a method of procedure less satisfactory, at least from the point of view of logic.

§ 34. Essential Equivalents of the Postulates.

If one is interested in the theory of relativity on account of its significance to physical science it is unnecessary to have complete logical equivalents of the postulates in order to justify

it. All that is essential is to find a set of postulates, experimentally demonstrable, by means of which it is possible to demonstrate the characteristic conclusions of relativity concerning the relations of units of time and units of length in two systems of reference.* Such a set of postulates we shall call essential equivalents of the postulates of relativity. The object of this section is to determine essential equivalents of postulate R, that is, such postulates as may be taken in connection with postulates M, V, L, so that the new set shall be essentially equivalent to M, V, L, R.

For this purpose let us first consider the relation between the transverse mass of a moving body and its mass at rest as given in theorem XI. Let us suppose that this theorem is true† (whether proved experimentally or otherwise); and let us seek its consequences. Suppose that the experiment by means of which we proved theorem XI is now repeated. If we again assume the law of conservation of momentum and equate the two observed changes in momenta, it is clear that we shall have a relation between measurements of time as carried out in the two systems of reference, and that this relation will be precisely the same as in the usual theory of relativity. Having this relation concerning time units we can then proceed as in the first paragraph in § 25 to derive the usual relations between units of lengths. Hence we have the following result:

THEOREM XVI. If m_0 and $t(m_0)$ have the same meaning as in theorem XI and if for any particular kind of matter whatever we have the relation

$$t(m_0) = \frac{m_0}{\sqrt{1-\beta^2}},$$

then this fact and postulates $(MVLC_1)$ form an essential equivalent of postulates $(MVLRC_1)$.

* It is obvious that we should then be able to demonstrate theorems XI and XII concerning the mass of a moving body.

† All that is essential to the argument is the truth of theorem XI for a particle of matter of some one kind; it need not be assumed to be true universally.

Next, let us suppose that for some particular kind of matter we have the relation

$$t(m_v) = (1 - \beta^2) l(m_v),$$

where $t(m_v)$ and $l(m_v)$ denote the transverse mass and the longitudinal mass, respectively, as in § 24. Then repeat the experiments by means of which we proved theorem XII. As before the balls will appear to B to move on arcs of the circle

$$x' = \cos\theta, \quad y' = \sin\theta.$$

Suppose that to A they appear to move along arcs of the ellipse*

$$x = \cos\theta, \quad y = \rho \sin\theta,$$

where ρ is a constant to be determined. As before, without the use of postulate R, it may be shown that to A the periods will be the same in the two cases. Then determine the periods as in the preceding discussion. The expression for the periods will contain ρ; in fact on equating them we shall find

$$m_2 = \rho^2 m_1.$$

But $m_1 = l(m_v)$ amd $m_2 = t(m_v)$; whence on account of the relation between $l(m_v)$ and $t(m_v)$, we have at once

$$\rho = \sqrt{1 - \beta^2}.$$

This, in connection with postulate L, leads readily to the usual relations concerning the units of length in two systems of reference. Having these relations of length units, the dimensional equation

$$V = \frac{L}{T}$$

taken in connection with postulate V leads at once to the usual relation of time units, provided we take the motion along the line of relative motion of the two systems. Hence we have the following theorem:

* Since we are assuming postulate L it is clear that the path must be of this form.

Theorem XVII. If $l(m_v)$ and $t(m_v)$ have the same meaning as in theorems XI and XII and if for any particular kind of matter whatever we have the relation

$$t(m_v) = (1 - \beta^2) \cdot l(m_v),$$

then this fact and postulates $(MVLC_1C_2)$ are essential equivalents of postulates $(MVLRC_1C_2)$.

§35. The Bucherer Experiment.

Our postulates V, L, C_1 have been universally accepted as part of the basis of the classical mechanics. Many persons have found no difficulty in accepting postulate M; certain it is at least that we have absolutely no evidence to contradict it. We have seen in theorem XVI that these four postulates, taken in connection with the formula for transverse mass, form an essential equivalent of $(MVLRC_1)$; in other words, the experimental demonstration of the formula for transverse mass carries with it the experimental proof of the theory of relativity, provided that postulates $(MVLC_1)$ are accepted as experimentally proved.

Bucherer (Annalen der Physik, ser. 4, vol. 28, pp. 513–536) has carried out some investigations which have been supposed to furnish this experimental verification for the formula of transverse mass, and hence for the whole theory of relativity. In order to draw this conclusion from Bucherer's direct results it is necessary to make use of a law which we have not yet employed, namely, the law of conservation of electricity which we have stated as postulate C_3. Since this law has customarily been accepted, we shall conclude that we have in Bucherer's results a partial experimental confirmation of the theory of relativity.

Bucherer's investigations have to do with the mass of a moving electron. There seems to be no means at hand for a direct measurement of this mass, and Bucherer resorted to the expedient of determining the ratio of charge to mass. Let us denote the charge by e, which we suppose to be constant,

in accordance with postulate C_3. As before let m_0 and $t(m_v)$ denote the mass at rest and the transverse mass when moving with velocity v, of the electron in consideration. Bucherer's experiments were carried out to determine the relation which exists between e/m_0 and $e/t(m_v)$. The measurements agreed in a remarkable way, not only as to general characteristics but also as to exact numerical results, with the formula*

$$\frac{e}{t(m_v)} = \frac{e\sqrt{1-\beta^2}}{m_0}.$$

Taking this formula as thus experimentally demonstrated we have at once our fundamental relation for transverse mass:

$$\sqrt{1-\beta^2}\cdot t(m_v) = m_0.$$

From this it follows that the experimental demonstration of the theory of relativity is complete when we have proved M, V, L, C_1 and C_3, provided that one accepts Bucherer's proof of the above relation between e/m_0 and $e/t(m_v)$. That is, *the essentials of the theory of relativity flow from principles for each of which there is strong experimental confirmation.* This important conclusion has often been pointed out.

To the present writer, however, it seems that one point especially should be subjected to further examination. Is it in fact true that the charge of a moving electron is independent of the velocity with which it moves? Let e_0 be the charge of the electron when at rest and denote by $t(e_v)$ its apparent charge when in motion with velocity v, the charge being measured by means of tests in which the line of action is perpendicular

* As a matter of fact Bucherer did not measure the ratio e/m_0. Instead of this he considered the ratio $e/t(m_v)$ for a considerable range of values for v and noticed that its value always agreed with the formula $e/t(m_v) = k\sqrt{1-\beta^2}$, where k is a constant. It appears natural, then, to assume that $m_0 = e/k$, whence one has the formula in the text. It should be emphasized that this assumption is necessary in order that the Bucherer results may be associated with our theorem as in the text, and consequently the conclusions there reached can be accepted with no stronger confidence than that which one has in the accuracy of the above assumption. See the next section where a possible means of experimental verification of the theory of relativity is suggested which does not depend on this assumption for its validity.

to the line of motion of the charge. In the above discussion we have assumed, in accordance with the usual practice, that $e_0 = t(e_v)$. Suppose however that the true relation were different from this, that, in fact, we have

$$t(e_v) = e_0\sqrt{1-\beta^2};$$

then Bucherer's experiment would lead to the conclusion that $t(m_v) = m_0$, and thus the whole theory of relativity would be overturned. Furthermore, if any relation other than $e_0 = t(e_v)$ is the true one, some modification at least of the theory of relativity would have to be made or else one would have to give up postulate C_1 which asserts the law of conservation of momentum. This result emphasizes the great importance of the question of the constancy of electric charge on the electron. We shall treat this matter further in the next section.

§ 36. Another Means for the Experimental Verification of the Theory of Relativity.

Just as theorem XVI was used for the theoretical basis of Bucherer's (partial) experimental demonstration of the theory of relativity so theorem XVII may be employed as the theoretical basis of a new experimental investigation which has not yet been carried out, one which bears the same essential relation as that of Bucherer to the confirmation or disproof of the entire theory of relativity. The object of this section is to indicate the nature of this experiment.

Let e_0 denote the charge of an electron when at rest with respect to a given system of reference. When it is in motion with a velocity v let $t(e_v)$ and $l(e_v)$ be the apparent charge when measured by means of tests whose lines of action are perpendicular and parallel, respectively, to the line of motion of the electron.

If we employ postulate C_3 we conclude that $e_0 = t(e_v) = l(e_v)$. We shall first assume the truth of one of these relations, namely, $t(e_v) = l(e_v)$, and we shall denote the common value of these two quantities by e. Now let us suppose that some means are

found for measuring both the quantities $e/t(m_v)$ and $e/l(m_v)$, where $t(m_v)$ and $l(m_v)$ denote as usual the transverse mass and the longitudinal mass respectively of the moving electron, whose velocity is v. Bucherer's methods furnish a means of measuring the first of these ratios; it will be necessary to devise a way to determine the value of the second ratio.

Or, instead of finding a means of measuring the two quantities $e/t(m_v)$ and $e/l(m_v)$ it will be sufficient if one determines only their ratio, as will be obvious from the discussion following.

Suppose now that we find the relation predicted by the theory of relativity:

$$\frac{e}{t(m_v)} = \frac{e}{(1-\beta^2)l(m_v)}.$$

This equation leads to the relation $t(m_v) = (1-\beta^2)\cdot l(m_v)$. According to theorem XVII this would give a new experimental confirmation of the theory of relativity. The importance of such a result is apparent.

But we should also have more than this. Having now concluded that the theory of relativity is confirmed and this result having been reached without the use of a relation between e_0 and $t(e_v)$, we may now use the experiment of Bucherer to draw further conclusions concerning electric charges in motion. In particular, it is obvious that we should have a proof of the fundamental relation

$$e_0 = t(e_v).$$

That is to say, having assumed that $t(e_v)$ and $l(e_v)$ are equal we conclude further on direct experimental evidence that each of these is equal to e_0. Now it is difficult to conceive how $t(e_v)$ and $l(e_v)$ could be different, for this would imply that the notion of electric charge is in need of essential modification. In fact, if the charged body is moving, the notion of charge would be indefinite in meaning until we had assigned the direction along which such charge is to be measured. Thus, if the experiment should turn out as surmised above, we should not only have the strongest sort of experimental confirmation of the theory of relativity but we should also have a valuable

verification of the fact that an electric charge does not vary in amount with the velocity of the body which carries it.

Suppose, on the other hand, that we make no assumption concerning the relation of $t(e_v)$ and $l(e_v)$ or of $t(m_v)$ and $l(m_v)$. On carrying out the experiments a relation of the form

$$\frac{t(e_v)}{t(m_v)} = k\frac{l(e_v)}{l(m_v)},$$

will be obtained where k is a constant or a variable depending on v. If it is found that k is different from unity we shall be forced to the conclusion that either our conception of mass in the classical mechanics or our conception of charge in the classical electrical theory is in need of essential modification. Again, if $k = 1$ and if we assume, as is natural, that $t(e_v) = l(e_v)$. then we have an experimental disproof of the theory of relativity. In fact we have such a disproof unless $k = 1/(1-\beta^2)$, provided of course that we assume $t(e_v) = l(e_v)$.

From these remarks it is obvious that, whatever may be the result of the experiments, they will certainly lead to important conclusions of a fundamental nature; that is, we have here a *crucial* experiment, one that cannot fail to lead somewhither. It is to be hoped that some laboratory worker will soon perform the requisite experiments; the writer, who is a mathematician, can only regret that he cannot conveniently carry out the work himself.

A variation of the experiment of Bucherer would seem to be sufficient for the purpose here. Bucherer's results were obtained by subjecting the moving electron to a magnetic field and also to an electric field each at right angles to the line of motion. A variation of the direction of these fields relative to the line of motion of the electron would probably afford a means of making the necessary measurements for the experimental proof of the relations requisite for use in the preceding discussion.

CHAPTER VII.

THE GENERALIZED THEORY OF RELATIVITY.

§ 37. Summary of Results from Previous Chapters.

We have seen (in § 20) that the restricted theory of relativity, as developed in the preceding pages, calls for a transformation of a remarkable kind between the coordinates of two systems moving with a uniform velocity v relatively to each other. If x, y, z, t are the space-time coordinates of a system S and x', y', z', t' are the space-time coordinates of a system S', related to S as in § 20, then we have

$$(1) \qquad t'=\gamma\left(t-\frac{v}{c^2}x\right), \quad x'=\gamma(x-vt), \quad y'=y, \quad z'=z,$$

where

$$\gamma=\frac{1}{\sqrt{1-\beta^2}}, \quad \beta=v/c,$$

and c is the velocity of light. If these equations are solved for x, y, z, t in terms of x', y', z', t', the resulting equations are what those in (1) become on interchanging the primed with the unprimed coordinates and replacing v by $-v$. The transformations thus have a remarkable symmetry.

From the first two equations in (1) we have

$$\frac{dx'}{dt'}=\frac{dx-vdt}{dt-vdx/c^2};$$

or,

$$(2) \qquad u'=\frac{u-v}{1-uv/c^2},$$

if $u'=dx'/dt'$ and $u=dx/dt$, these denoting the velocities of motion in the x-direction of the coordinate axes in each system

with respect to that system. Formula (2) states the principle of addition of velocities, from it, as we have seen (in § 21), it follows that the usual law of addition of velocities is not maintained.

But Robb has pointed out that if we introduce the notion of *rapidity* of motion and define the measure of rapidity of motion with a velocity v to be $\tanh^{-1}(v/c)$, then we do have for rapidities a simple law of addition, namely, that obtained readily from equation (2) and expressible in the form

$$\tanh^{-1}\left(\frac{u'}{c}\right) = \tanh^{-1}\left(\frac{u}{c}\right) - \tanh^{-1}\left(\frac{v}{c}\right).$$

Again, we have found (in § 24) that the mass of a body is dependent upon its velocity relative to the observer's system, and in fact that the transverse mass $t(m_v)$ of a body in motion with a uniform velocity v relative to the system is expressed in terms of the mass m_0 of the body at rest by the formula

$$t(m_v) = \frac{m_0}{\sqrt{1-\beta^2}}.$$

The formula for longitudinal mass is also given in § 24; since we have no further need for the conception of longitudinal mass that formula will not be repeated here. Hereafter we shall use the word "mass" to denote what we have heretofore called transverse mass.

In § 26 we have seen that mass and energy are interchangeable in the sense that the measure of each may be expressed directly in terms of the measure of the other. A consequence of this is that we may use in our equations energy and not mass or mass and not energy instead of both mass and energy, if it should turn out that such a thing shall serve our convenience.

In the generalized theory of relativity it can hardly be said that the notion of force enters at all; consequently we shall not need to employ in this chapter the formula (of § 25) for the transformation of force in passing from one system of reference to another.

§ 38. Transformations in Space of Four Dimensions.

Following a suggestion of Minkowski's, we may look upon the transformation (1) as in a certain sense the transformation due to a rotation of axes in a space-time extension of four dimensions. For the purpose of viewing it in this light let us assume that the units are so chosen that the velocity c of light is unity and let us make the imaginary transformation

$$t=i\tau,\quad t'=i\tau',\quad v=i\tan\theta,\quad i=\sqrt{-1}.$$

Then we have $\gamma=\cos\theta$ and equations (1) reduce readily to the form

$$(3)\qquad \tau'=\tau\cos\theta-x\sin\theta,\quad x'=x\cos\theta+\tau\sin\theta,\quad y'=y,\quad z'=z.$$

This represents a rotation of the axes of coordinates through an imaginary angle θ in the $x\tau$-plane.

In the Newtonian mechanics the laws of nature are assumed to be invariant with respect to a change in the orientation of axes in the xyz-space. In the theory of relativity this principle is extended to the four-dimensional space-time $xyz\tau$-extension; and it is therefore assumed that this four-dimensional extension is completely isotropic. From this point of view the conspiracy of nature to prevent our measuring the velocity of our system through space disappears; there is nothing to conceal. Space-time extension being isotropic, there is no variation of properties in different directions and hence nothing for us to detect; no one orientation is more fundamental than another.

One consequence of this is that we cannot pick out one direction as the absolute time any more than we can pick out one direction as the absolute vertical. As Minkowski has said: "Henceforth space and time in themselves vanish to shadows, and only a kind of union of the two preserves an independent existence." Unfortunately the simplicity of the conception of a four-dimensional space-time extension and the interpretation of our transformation (1) as due to a rotation in it are marred

by the fact that nature makes a distinction of such sort that we have to employ the imaginary time variable τ in order to exhibit our transformation as having the properties of a rotation in a space of four dimensions.

If we use ids to represent the element of "length" in our space-time extension, that is, the interval between two point-events, its value will be given by the relation

$$(4) \qquad -ds^2 = dx^2 + dy^2 + dz^2 + d\tau^2.$$

It is easy to see that ds is invariant with respect to a rotation of axes and in particular with respect to the rotation defined by equations (3). If we choose axes moving with the particle we have, $dx = dy = dz = 0$, so that $ds^2 = -d\tau^2$ or $ds = dt$; this explains the choice of sign in the first member of (4), the purpose being to secure in s a sort of time variable.

It can be shown that the whole restricted theory of relativity, as developed in the preceding chapters, is summed up in the conclusion that ds is invariant; and from the hypothesis of the invariance of ds one can deduce the transformation equations (1), and hence the other fundamental results of the theory.

We have used the $xyz\tau$-extension for convenience in deriving certain geometric properties of the transformation (1). It is clear that we may likewise look upon x, y, z, t as coordinates of a point in the real space-time $xyzt$-extension of four dimensions. Let us consider a little more closely this space-time extension. Each point P of this four-dimensional extension represents both a definite place A in the usual space of three dimensions and a definite moment of time t at which this place A is to be considered. If P refers to a material point it shows the time t at which this point is found at the place A. In the course of time the material point is represented every moment by a new point P of the space-time extension; all these points P lie on a world line which represents completely once for all the state of motion or of rest of the material point for all time. In the same way we may speak of the world-line of any event in nature. An intersection of two world-lines indicates that the two objects to which they belong meet at a certain moment,

that a "coincidence" takes place. Similarly, one may represent statically in a space of three dimensions the kinetics of a body moving in a plane; a clear picture of this special case will assist one greatly in grasping the more general considerations for motion in a space of three dimensions.

Now Einstein has remarked that the only things which we can observe and measure among events in nature are these coincidences, the intersections of these world-lines, and that it is with these alone that our theories are essentially concerned. From this it follows that the results of all observations may be represented by world-lines in a four-dimensional extension—let us say by means of a field-figure—and that the only things directly observed are the intersections of these lines one with another.

In our statement of the laws of nature we shall have to attend only to the intersections of the world-lines; and hence we shall have a great liberty in the construction of the field-figures. We may vary this construction in any way we please, provided only that we do not disturb the order of the intersections of the world-lines. As the number of observed intersections increases we are more restricted in this freedom, but only in the way of a finite number of added restrictions to the infinitude of possible changes inherent in the nature of the process. Even if all the intersections in nature were known there would still be great freedom in the construction of the field-figures. If two persons independently represent the same observations by means of world-lines their field-figures will probably be quite different and in fact will probably agree only as to the order of the points of intersection, all other properties of the two figures being different; and indeed these figures will quite as well represent the observed facts after being deformed in any way provided that there is no break in continuity during the process of deformation.

One consequence of these considerations, which is important for our purposes, is that there is a great freedom of choice of coordinate systems in reducing the observed laws to analytical form and that the essential laws may be represented quite as

well on the basis of one of these systems as on that of another. For a long time it has been customary to introduce various types of curvilinear coordinates or of moving axes for the study of particular problems in physics, but it has usually not been forgotten even for a moment that these coordinate systems were curvilinear or were in motion. But there is at least one important and instructive case in which a simple means has been found for ignoring the peculiarity of the axes during the mathematical investigation. This is the case of rotating axes.

In many dynamical systems, some part of the system is compelled to rotate with constant angular velocity ω round a given fixed axis. The motion of a bead on a rotating twisted wire furnishes a simple example. The system might be treated by a direct application of Lagrange's equations; but it is often more convenient to use a principle which reduces the consideration of systems of this kind to that of systems in which no rotation takes place. When one develops the general differential equations of motion of such a system (as in Whittaker's *Analytical Dynamics*, second edition, pp. 40–41, for instance) it is seen that the motion is the same as if the prescribed angular velocity were zero and the potential energy contained an additional term $-\frac{1}{2}\Sigma mr^2\omega^2$, where m denotes the mass and r the distance of a particle from the axis of rotation. Thus, by modifying the potential energy, we may replace the consideration of a system which is constrained to rotate uniformly about an axis by that of a system for which this rotation does not take place. The imaginary forces which are introduced in this way to represent the acceleration effect of the enforced rotation are often called *centrifugal forces*.

Now let us suppose that an observer is stationed on such a rotating system and that he is shut off from the observation of things external to his system in such a way that he is quite unaware of the fact that it is rotating. The fictitious term which we have just supposed to be added to the potential energy to account mathematically for the rotation would not be fictitious for him but would represent an existing part of the actual potential. He would not unnaturally take it to be due

to a part of his gravitational field. At any rate he would have no means to distinguish it from a potential due to a changing component of the gravitational field. To him then this centrifugal force would seem to be a real thing; but to us, who look upon his system from the outside, this centrifugal force appears to be fictitious.

Now suppose that a ray of light passes across this observer's rotating system. We who look upon his system from the outside will see that the ray passes in a straight line. But to him, on account of the unobserved rotation of his system, it will appear to move on a curved path. If he sets out to investigate the curvature of this path he will find that it depends upon that element of his gravitational field which we consider to be fictitious and he will conclude that the ray of light is bent in its course by the gravitational field, or at least by a certain part of it.

We shall not pursue this further in the loose way of this section but shall leave it to be taken up again more rigorously when we have prepared suitable mathematical machinery for dealing with it.

§ 39. The Principle of Equivalence.

There are no coordinate axes in nature; these are introduced by us for convenience in the analytical representation of phenomena. When we enunciate the laws of mechanics and electrodynamics with respect to "unaccelerated rectangular axes," or "Galilean axes" as they are sometimes called, the only definition which we can give of these axes is that they are the axes with respect to which the laws may be enunciated correctly in the form in which we state them. We cannot recognize such axes intuitively. One fundamental and central purpose of the generalized theory of relativity is to restate the laws of nature in such a form that the statement shall be independent of the system of coordinates and hence be equally applicable to all systems.

When we introduced the centrifugal force in the preceding section and then looked upon the rotating system of reference as stationary so that moving bodies were acted upon by a

fictitious force, we saw that the changed point of view gave us a space in which all paths (even that of light) were deformed in a purely geometric way. Everything was acted upon in the same way by this fictitious force. This property is also shared by the force of gravitation; the gravitational field produces an acceleration which is independent of the nature or the mass of the body acted upon. This has led to the hypothesis that the force of gravitation may be, so far as we can observe, of essentially the same nature as the centrifugal or geometrical forces introduced by the choice of coordinates.

This hypothesis has been framed by Einstein into his now celebrated Principle of Equivalence. It may be enunciated as follows: A gravitational field of force is exactly equivalent to a field of force introduced by a transformation of the coordinates of reference so that we cannot by any possible experiment distinguish between them.

Eddington in his report to the Physical Society of London on "The Relativity Theory of Gravitation" has insisted on a precise criterion for the cases in which the principle of equivalence is *assumed* to apply. He formulates it as follows: The laws, relating to phenomena in a geometric field of force, which depend on the coefficients g of the next section and their first derivatives, will also hold in a permanent gravitational field, namely, one that cannot be entirely removed merely by a change of axes; but laws which depend on the second or higher derivatives of the g's will not necessarily have this universality.

It will be observed that this contention rests upon an implied limitation of the principle of equivalence. It is argued that there is such a thing as a natural or permanent gravitational field of force which cannot be altogether transformed away. There is a tacit agreement that a natural gravitational field of force exists even though it may be altogether impossible for us to distinguish between it and the fictitious geometrical or centrifugal forces due to the choice of coordinate axes.

Transformations exist which remove the gravitational field at a point; but we cannot find any transformation which will remove the gravitational field throughout a finite region.

Since we cannot distinguish between the fictitious and the permanent forces we shall have no means in general of selecting any particular system of coordinates as fundamental. We must therefore state the laws of nature in a form which is quite independent of the choice of axes.

§ 40. General Transformation of Axes.

If we seek a suitable analytical expression of the freedom which we have seen to exist in our construction of the field-figures of natural phenomena we shall be led to certain considerations which are of essential importance for our purposes. If we represent in terms of a given set x, y, z, t of space-time coordinates the configuration of a system of world-lines and in terms of a set x', y', z', t' the configuration of another system of world-lines, it is not difficult to see that the two systems of world-lines will have corresponding intersections and these arranged in the same relative order when and only when there exists a single-valued continuous transformation from the coordinates of each system to those of the other.

Since events do not presuppose any particular system of coordinates and the space-time scaffolding is introduced by us merely for our convenience, it is desirable to have our laws of nature expressed so as to be quite independent of the system of coordinates employed. Let us see what this amounts to in the case of the element ds whose value in the absence of a gravitational field is given, as we have seen, by the equation,

$$ds^2 = -dx^2 - dy^2 - dz^2 + dt^2. \tag{5}$$

Let us introduce new coordinates x_1, x_2, x_3, x_4 by means of transformations of the kind just mentioned; they are given by equations of the form,

$$x = f_1(x_1, x_2, x_3, x_4), \quad y = f_2(x_1, x_2, x_3, x_4), \text{ etc.}$$

Then we have

$$dx = \frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2 + \frac{\partial f_1}{\partial x_3} dx_3 + \frac{\partial f_1}{\partial x_4} dx_4, \text{ etc.} \tag{6}$$

Putting such values of dx, dy, dz, dt into equation (5), we have a relation which may be written in the form

$$(7)\qquad ds^2 = g_{11}dx_1^2 + g_{22}dx_2^2 + g_{33}dx_3^2 + g_{44}dx_4^2 + 2g_{12}dx_1dx_2 + 2g_{13}dx_1dx_3 + 2g_{14}dx_1dx_4 + 2g_{23}dx_2dx_3 + 2g_{24}dx_2dx_4 + 2g_{34}dx_3dx_4,$$

where the g's are readily computed functions of the coordinates x_1, x_2, x_3, x_4, depending on the functions f_1, f_2, f_3, f_4 of the transformation. We shall henceforth assume that ds is defined by an equation of the form (7), without reference to whether that equation is derived from (5) or by other means.

If we take the particular transformation of rotating axes

$$x = x_1 \cos \omega x_4 - x_2 \sin \omega x_4, \quad y = x_1 \sin \omega x_4 + x_2 \cos \omega x_4, \quad z = x_3, \quad t = x_4,$$

we obtain readily the relation.

$$(8)\qquad ds^2 = -dx_1^2 - dx_2^2 - dx_3^2 + [1 - \omega^2(x_1^2 + x_2^2)]dx_4^2 + 2\omega x_2 dx_1 dx_4 - 2\omega x_1 dx_2 dx_4.$$

By comparing this with (7) we obtain the values of the g's for this system of coordinates. In particular we have,

$$g_{44} = 1 - 2\Omega, \qquad \Omega = \tfrac{1}{2}\omega^2(x_1^2 + x_2^2),$$

where Ω is the potential of the centrifugal force.

From this it follows that the coefficient g_{44} may be regarded as a potential; and this conception is extended so that all the the coefficients g are regarded as components of a generalized potential of the field of force. It is unnecessary and indeed experimentally impossible, according to the theory of relativity, to distingush between the portion of the g's arising from the choice of coordinates and that arising from the so-called natural or permanent gravitational field. We shall usually speak of the entire field as gravitational and shall say that a gravitational field is specified by a set of values of the g's whatever their source.

These coefficients g may be looked upon in two ways. In one aspect they may be thought of as expressing the metrical properties of the coordinates; and this is the orthodox standpoint of the theory of relativity; it is attained by banishing

almost or quite entirely the notion of gravitational force, this exclusion being based on the fact that if there are such things as actual gravitational forces we cannot distinguish them from fictitious geometrical forces.

The values of the g's given by equation (5), namely, $g_{11}=-1$, $g_{22}=-1$, $g_{33}=-1$, $g_{44}=1$, $g_{ij}=0$, when $i \neq j$, are those which obtain in the absence of a gravitational field and for a suitable choice of reference system. We may call them the *Galilean values of the* g's. When the coordinates can be chosen so that the g's have these values we may regard coordinates so chosen as fundamental; and the deviations of the g's for any other choice of coordinates may be looked upon as due to the distortion of the space-time extension or to the gravitational field.

There is a general limitation imposed on the g's not by mathematics but by nature; and this limitation is expressible by means of differential equations satisfied by the g's and thus exhibiting the law of gravitation; these differential equations we shall find later.

Now the g's vary with the system of coordinates employed. But the essential law is independent of the system of coordinates. If new coordinates are chosen we get new values of the g's through use of the transformed form of (7) and the hypothesis of the invariance of ds. The differential equations between the new g's and their coordinate system must be the same as those between the old g's and their coordinate system. In other words, the differential equations expressing the law of gravitation must be covariant under the general transformations of axes which we have defined. This fact will furnish us with a valuable guide in the development of the new laws.

Moreover, if we have equations expressing some physical law in the usual coordinates and we are able to recognize these equations as the degenerate form of equations covariant under our general group of transformations we may hope to be able to rewrite the known equations in a more far-reaching form,

so that they shall then afford the extended law on the basis of the theory of relativity. It is by this method of approach indeed that we shall obtain the differential equations which characterize the gravitational field.

§ 41. The Theory of Tensors.

In order to proceed with our problem of finding the differential equations of the general gravitational field we shall need the properties of certain fundamental mathematical symbols; these we treat in this and the two following sections in so far as they are needed for our present purposes.

We shall be concerned particularly with differential equations which are covariant under the group of continuous transformations,

$$(9) \qquad x_i' = f_i(x_1, x_2, x_3, x_4), \quad i = 1, 2, 3, 4.$$

This is due to the fact that in the general theory of relativity the laws of nature are to be expressed in a form which is covariant under the transformations (9). Here we assume that the functions f_i are of such sort that the Jacobian

$$J = \left| \frac{\partial x_i}{\partial x_j} \right|$$

of the transformation is different from zero throughout the entire range of values in consideration. That these transformations form a group follows from the existence of a unique inverse of the transformation (due to the non-vanishing of the Jacobian) and the fact that the product of two transformations of the set also belongs to the set.

Now in view of (9) we have

$$dx_i' = \Sigma_j \frac{\partial x_i'}{\partial x_j} dx_j, \qquad \frac{\partial \phi}{\partial x_i'} = \Sigma_j \frac{\partial x_j}{\partial x_i'} \frac{\partial \phi}{\partial x_j}, \qquad i = 1, 2, 3, 4,$$

where ϕ is a function of x_1, x_2, x_3, x_4 and Σ_j denotes the sum as to j for $j = 1, 2, 3, 4$. (We shall employ similarly the symbols $\Sigma_{\mu\nu}$, $\Sigma_{\alpha\beta\gamma}$, etc., to denote the sum of the elements formed

from the term following the symbol by giving independently to the suffixes μ, ν or α, β, γ, etc., the values 1, 2, 3, 4.) We may look upon these equations as transforming the vectors (dx_1, dx_2, dx_3, dx_4) and ($\partial\phi/\partial x_1, \ldots, \partial\phi/\partial x_4$) into the vectors ($dx_1', dx_2', dx_3', dx_4'$) and ($\partial\phi/\partial x_1', \ldots, \partial\phi/\partial x_4'$), respectively.

If two vectors A^μ and A_μ are transformed through (9) by the first and second of these laws, respectively, so that we have

$$A'^\mu = \Sigma_\sigma \frac{\partial x_\mu'}{\partial x_\sigma} A^\sigma, \qquad A_\mu' = \Sigma_\sigma \frac{\partial x_\sigma}{\partial x_\mu'} A_\sigma, \tag{10}$$

we shall say that A^μ is a *contravariant* vector and that A_μ is a *covariant* vector. Moreover, we shall uniformly mean by the symbols A^μ, B^μ, ... contravariant vectors; and by the symbols A_μ, B_μ, ... covariant vectors.

If μ and ν each ranges over the set 1, 2, 3, 4, the symbol $A_{\mu\nu}$ will denote a quantity having sixteen components. Similarly the symbols $A_{\mu\nu\sigma}$ and $A^\sigma_{\mu\nu}$ will each denote a quantity having sixty-four components. We may look upon such quantities as a sort of generalized vectors. Since the present theory is concerned especially with those for whose components the transformation equations are linear and homogeneous it is found convenient to apply a particular name to such vectors; and they are called *tensors*. If then a law of nature is so formulated as to be expressed through the vanishing of all the components of a tensor, it will be covariant under the transformations (9). Tensors are therefore of central inportance for the theory of relativity.

By an extension of the terminology employed in connection with (10) we may speak of covariant, contravariant and mixed tensors; those for two indices (or those of rank two) obey by definition the transformation laws:

$$A'_{\mu\nu} = \Sigma_{\sigma\tau} \frac{\partial x_\sigma}{\partial x_\mu'} \frac{\partial x_\tau}{\partial x_\nu'} A_{\sigma\tau} \quad \text{(covariant tensor)}, \tag{11}$$

$$A'^{\mu\nu} = \Sigma_{\sigma\tau} \frac{\partial x_\mu'}{\partial x_\sigma} \frac{\partial x_\nu'}{\partial x_\tau} A^{\sigma\tau} \text{(contravariant tensor)}, \tag{12}$$

$$(13) \qquad A_\mu'^{\nu} = \Sigma_{\sigma\tau} \frac{\partial x_\sigma}{\partial x_\mu'} \frac{\partial x_\nu'}{\partial x_\tau} A_\sigma^\tau \qquad \text{(mixed tensor).}$$

It will be observed that the notation for each type of tensor is so chosen as to indicate the character of the tensor. Tensors of the third and higher rank are so defined that analogous laws of transformation obtain; thus for covariant tensors of the third rank we have the law

$$A'_{\lambda\mu\nu} = \Sigma_{\rho\sigma\tau} \frac{\partial x_\rho}{\partial x_\lambda'} \frac{\partial x_\sigma}{\partial x_\mu'} \frac{\partial x_\tau}{\partial x_\nu'} A_{\rho\sigma\tau}.$$

Vectors such as those involved in (10) may be called tensors of rank unity. A scalar (invariant) may be called a tensor of rank zero and classed as either covariant or contravariant.

If we now introduce a third set x_λ'' of coordinates as functions of the set x_λ' and if A_λ'' is the transformed vector of A_λ' when transformed by this substitution, we have

$$A_\lambda'' = \Sigma_\mu \frac{\partial x_\mu'}{\partial x_\lambda''} A_\mu', \qquad A_\mu' = \Sigma_\sigma \frac{\partial x_\sigma}{\partial x_\mu'} A_\sigma.$$

But

$$\Sigma_\mu \frac{\partial x_\sigma}{\partial x_\mu'} \frac{\partial x_\mu'}{\partial x_\lambda''} = \frac{\partial x_\sigma}{\partial x_\lambda''}; \qquad \text{hence} \qquad A_\lambda'' = \Sigma_\sigma \frac{\partial x_\sigma}{\partial x_\lambda''} A_\sigma.$$

From this it follows that the final result obtained by applying the two transformations successively to A_σ is the same as that obtained by applying at once the product of the two transformations. It is not difficult to see that this transitive property is possessed by the tensors of each of the several classes.

It is evident that the sum of any two tensors of the same given character is also a tensor of that character.

It may also be proved that the product of two tensors is a tensor and that its character is the sum of the characters of the two component tensors. Let us prove this for the single case of a product of the form $A_{\mu\nu} B_\lambda^\rho$. From the relations of transformation

$$A'_{\mu\nu} = \Sigma_{\alpha\beta} \frac{\partial x_\alpha}{\partial x_\mu'} \frac{\partial x_\beta}{\partial x_\nu'} A_{\alpha\beta}, \qquad B_\lambda'^\rho = \Sigma_{\gamma\delta} \frac{\partial x_\gamma}{\partial x_\lambda'} \frac{\partial x_\rho'}{\partial x_\delta} B_\gamma^\delta,$$

we have

$$(A'_{\mu\nu}B'^{\rho}_{\lambda}) = \Sigma_{\alpha\beta\gamma\delta}\frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial x_\beta}{\partial x_\nu'}\frac{\partial x_\gamma}{\partial x_\lambda'}\frac{\partial x_\rho'}{\partial x_\delta}(A_{\alpha\beta}B_\gamma{}^\delta).$$

Hence the law of transformation is that of a tensor of the fourth rank having the character denoted by the symbol $C^{\rho}_{\mu\nu\lambda}$.

In particular, the product of two vectors is a tensor of the second rank; there are also tensors of the second rank which are not products of two vectors.

In the two preceding paragraphs we used the term product of two tensors to denote the tensor whose component elements are all the elements formed by multiplying an element of one tensor by an element of another tensor. This may be called their *outer product*. We need also the notion of *inner product* of two vectors, say of A_μ and B^μ; and this is defined to be the quantity $\Sigma_\mu A_\mu B^\mu$; that is, the sum of products of corresponding elements.

From a mixed tensor such as $A^{\gamma}_{\mu\nu\sigma}$ we can form a contracted tensor $\Sigma_\sigma A^{\sigma}_{\mu\nu\sigma}$ which we denote by $B_{\mu\nu}$. Let us prove that the last quantity is indeed a tensor. We have

$$B'_{\mu\nu} = \Sigma_\sigma A'^{\sigma}_{\mu\nu\sigma}$$

$$= \Sigma_{\sigma\alpha\beta\gamma\delta}\frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial x_\beta}{\partial x_\nu'}\frac{\partial x_\gamma}{\partial x_\sigma'}\frac{\partial x_\sigma'}{\partial x_\delta}A^{\delta}_{\alpha\beta\gamma}.$$

But

$$\Sigma_\sigma\frac{\partial x_\gamma}{\partial x_\sigma'}\frac{\partial x_\sigma'}{\partial x_\delta} = \frac{\partial x_\gamma}{\partial x_\delta} = \begin{cases} 0 \text{ if } \gamma \neq \delta \\ 1 \text{ if } \gamma = \delta. \end{cases}$$

Hence,

$$\Sigma_{\sigma\gamma\delta}\frac{\partial x_\gamma}{\partial x_\sigma'}\frac{\partial x_\sigma'}{\partial x_\delta}A^{\delta}_{\alpha\beta\gamma} = \Sigma_\gamma A^{\gamma}_{\alpha\beta\gamma} = B_{\alpha\beta}.$$

Therefore,

$$B'_{\mu\nu} = \Sigma_{\alpha\beta}\frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial x_\beta}{\partial x_\nu'}B_{\alpha\beta},$$

showing that $B_{\mu\nu}$ is a tensor of the second rank of the character already anticipated in the notation. This process of contraction is applied to remove from the symbol any two indices of different character but never two of the same character.

To prove that a given quantity is a tensor of given character we may merely verify, as in the preceding paragraph, that its equations of transformation are those by which the tensor character is defined. But the result may often be more readily obtained by the use of the following theorem: *If the inner product of a given quantity by every covariant (or by every contravariant) vector is a tensor then the given quantity is itself a tensor.*

The general method of argument will be seen from the following special case: Suppose that $\Sigma_\nu A_{\mu\nu}B^\nu$ is a covariant vector for every choice of the contravariant vector B^ν. Then we have

$$\Sigma_\nu A'_{\mu\nu}B'_\nu = \Sigma_{\sigma\tau}\frac{\partial x_\sigma}{\partial x_\mu'}A_{\sigma\tau}B^\tau, \qquad B^\tau = \Sigma_\nu\frac{\partial x_\tau}{\partial x_\nu'}B'^\nu,$$

the last relation coming from the inverse transformation. Therefore,

$$\Sigma_\nu B'^\nu\left(A'_{\mu\nu} - \Sigma_{\sigma\tau}\frac{\partial x_\sigma}{\partial x_\mu'}\frac{\partial x_\tau}{\partial x_\nu'}A_{\sigma\tau}\right) = 0.$$

Since B'^ν is arbitrary it follows that the parentheses quantity must have the value zero for every value of μ and ν, showing that $A_{\mu\nu}$ is a covariant tensor of rank two. A similar proof can evidently be made for tensors of any given character.

Now in accordance with the theory of relativity the general expression for ds^2,

$$ds^2 = \Sigma_{\mu\nu}g_{\mu\nu}dx_\mu dx_\nu, \qquad g_{\mu\nu} = g_{\nu\mu},$$

is to be invariant under our transformations. Here dx_μ plays the role of an arbitrary contravariant vector, whence it follows by use of the preceding theorem that $\Sigma_\nu g_{\mu\nu}dx_\nu$ is a covariant tensor of the first rank. Repeating the argument with respect to this tensor we see that $g_{\mu\nu}$ is a covariant tensor of the second rank. We call it the *fundamental covariant tensor* on account of its central position in the theory of gravitation.

If in the determinant $g = |g_{\mu\nu}|$ we take the cofactor of the element $g_{\mu\nu}$ and divide this by the determinant g we obtain

certain quantities $g^{\mu\nu}(g^{\mu\nu}=g^{\nu\mu})$; these define a contravariant tensor, as we shall now show. We call it the *fundamental contravariant tensor*. From the theory of determinants we have $\Sigma_\sigma g_{\mu\sigma}g^{\nu\sigma}=g_\mu{}^\nu$, where $g_\mu{}^\nu$ is 1 or 0 according as μ and ν are equal or unequal. Since $\Sigma_\sigma g_\sigma{}^\nu A^\sigma=A^\nu$ for every contravariant vector A^ν, it follows that $g_\mu{}^\nu$ is a mixed tensor of the character anticipated by the notation. We call it the *fundamental mixed tensor*. We now have readily

$$ds^2=\Sigma_{\mu\nu\sigma}g_{\mu\sigma}g_\nu{}^\sigma dx_\mu dx_\nu=\Sigma_{\mu\nu\sigma\tau}g_{\mu\sigma}g_{\nu\tau}g^{\sigma\tau}dx_\mu dx_\nu.$$

Introducing the notation $d\xi_\sigma=\Sigma_\mu g_{\mu\sigma}dx_\mu$, we have

$$ds^2=\Sigma_{\sigma\tau}g^{\sigma\tau}d\xi_\sigma d\xi_\tau.$$

Since $d\xi_\sigma$ is an arbitrary covariant vector it follows from the last relation and the fact that $g^{\sigma\tau}=g^{\tau\sigma}$ that $g^{\sigma\tau}$ is a contravariant tensor of the character indicated.

With any covariant tensor $A_{\mu\nu}$ we may have the following two *associated tensors* and *scalar* of character indicated by the notation:

$$(14)\qquad A_\mu{}^\nu=\Sigma_\alpha g^{\nu\alpha}A_{\mu\alpha},\qquad A^{\mu\nu}=\Sigma_\alpha g^{\mu\alpha}A_\alpha{}^\nu,\qquad A=\Sigma_\mu A_\mu{}^\mu.$$

If $g'_{\mu\nu}$ are the quantities into which the $g_{\mu\nu}$ transform through (9) and if g and g' denote the determinants $|g_{\mu\nu}|$ and $|g'_{\mu\nu}|$ while J is the Jacobian of the transformation we have readily from the theory of determinants that

$$J^2g'=g$$

If $d\tau$ and $d\tau'$ are the elements of four-dimensional volume in our space-time extension we have $d\tau'=Jd\tau$, so that from the foregoing relation it follows that

$$\sqrt{-g}\,d\tau=\sqrt{-g'}\,d\tau'.$$

Now if we employ our hypothesis that in infinitely small regions the special relativity theory is valid we see that coordinates may be chosen so that ds^2 has the form given in (5), the

corresponding determinant g having the value -1. If $d\tau_0$ denotes the volume element in this system, the "natural" volume element, we have $d\tau_0 = \sqrt{-g}\,d\tau$. We see that g cannot vanish since then we should have for a finite volume element $d\tau$ an infinitesimal volume element $d\tau_0$. We shall assume that g has always a finite negative value, this assumption being in agreement with the special theory of relativity.

Since $-g$ is always positive and finite there must exist a set of axes for which it has the value 1. Expressed in terms of such a set of axes the laws of nature will have a particularly simple form. One should first derive them in their general covariant form and afterwards simplify them by the special choice of axes indicated, the simplification being effected merely as a matter of convenience.

§ 42. Covariant Differentiation.

The writing of certain important covariant differential expressions is greatly facilitated by the use of Christoffel's 3-index symbols, namely,

$$[\mu\nu, \lambda] = \tfrac{1}{2}\left(\frac{\partial g_{\mu\lambda}}{\partial x_\nu} + \frac{\partial g_{\nu\lambda}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\lambda}\right),$$

$$\{\mu\nu, \lambda\} = \tfrac{1}{2}\Sigma_\alpha g^{\lambda\alpha}\left(\frac{\partial g_{\mu\alpha}}{\partial x_\nu} + \frac{\partial g_{\nu\alpha}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\alpha}\right).$$

These symbols satisfy the relations

$$(15)\qquad \{\mu\nu, \lambda\} = \Sigma_\alpha g^{\lambda\alpha}[\mu\nu, \alpha], \qquad [\mu\nu, \lambda] = \Sigma_\alpha g_{\lambda\alpha}\{\mu\nu, \alpha\}.$$

By $[\mu\nu, \lambda]'$ and $\{\mu\nu, \lambda\}'$ we denote the quantities obtained from $[\mu\nu, \lambda]$ and $\{\mu\nu, \lambda\}$ respectively on replacing x_μ by x_μ' and $g_{\mu\nu}$ by $g'_{\mu\nu}$.

If in the relation

$$g'_{\mu\nu} = \Sigma_{\alpha\beta}\frac{\partial x_\alpha}{\partial x_\mu'}\frac{\partial x_\beta}{\partial x_\nu'}g_{\alpha\beta}$$

one differentiates with respect to x_λ', x_μ', x_ν', and subtracts member by member the first resulting relation from the sum

member by member of the last two, one has a result which is readily put in the form

$$[\mu\nu, \lambda]' = \Sigma_{\alpha\beta} g_{\alpha\beta} \frac{\partial^2 x_\alpha}{\partial x_\mu' \partial x_\nu'} \frac{\partial x_\beta}{\partial x_\lambda'} + \Sigma_{\alpha\beta\gamma} \frac{\partial x_\alpha}{\partial x_\mu'} \frac{\partial x_\beta}{\partial x_\nu'} \frac{\partial x_\gamma}{\partial x_\lambda'} [\alpha\beta, \gamma].$$

Multiplying by $g'^{\lambda\rho} \partial x_\epsilon / \partial x_\rho'$, summing as to λ and ρ, and simplifying by use of relations (12), (15), and properties of the g's given in the paragraph ending with (14), we have

$$(16) \qquad \Sigma_\rho \{\mu\nu, \rho\}' \frac{\partial x_\epsilon}{\partial x_\rho'} = \frac{\partial^2 x_\epsilon}{\partial x_\mu' \partial x_\nu'} + \Sigma_{\alpha\beta} \frac{\partial x_\alpha}{\partial x_\mu'} \frac{\partial x_\beta}{\partial x_\nu'} \{\alpha\beta, \epsilon\}.$$

On differentiating a scalar quantity one obtains a covariant tensor of rank unity; but on differentiating a tensor of rank greater than zero one obtains a quantity which is not necessarily a tensor. It is therefore desirable to define a process generalizing that of differentiation and of such sort as to lead always from a given tensor to a new tensor. Such a process we now define.

If we differentiate with respect to x_ν' both members of the second equation in (10) and in the result replace the second derivatives by their values taken from (16) we have

$$(17) \qquad \frac{\partial A_\mu'}{\partial x_\nu'} - \Sigma_{\rho\sigma} \{\mu\nu, \rho\}' A_\sigma \frac{\partial x_\sigma}{\partial x_\rho'} = \Sigma_{\sigma\tau} \frac{\partial x_\sigma}{\partial x_\mu'} \frac{\partial x_\tau}{\partial x_\nu'} \frac{\partial A_\sigma}{\partial x_\tau}$$
$$- \Sigma_{\alpha\beta\sigma} \frac{\partial x_\alpha}{\partial x_\mu'} \frac{\partial x_\beta}{\partial x_\nu'} \{\alpha\beta, \sigma\} A_\sigma.$$

Simplifying the second term of the first member by means of the second relation in (10) and introducing the symbol $A_{\mu\nu}$ with the meaning

$$(18) \qquad A_{\mu\nu} = \frac{\partial A_\mu}{\partial x_\nu} - \Sigma_\rho \{\mu\nu, \rho\} A_\rho,$$

we see from (17) that $A_{\mu\nu}$ is transformed in accordance with relations (11) so that $A_{\mu\nu}$ is a tensor of the character indicated by its symbol. This tensor $A_{\mu\nu}$ is called the *covariant derivative* of A_μ.

In a somewhat similar manner one can obtain formulæ for

defining the *covariant derivatives* of contravariant and mixed tensors, namely,

$$(19) \qquad A_\nu^\mu = \frac{\partial A^\mu}{\partial x_\nu} + \Sigma_\epsilon \{\nu\epsilon, \mu\} A^\epsilon,$$

$$(20) \qquad A_\nu^{\lambda\mu} = \frac{\partial A^{\lambda\mu}}{\partial x_\nu} + \Sigma_\epsilon \{\nu\epsilon, \lambda\} A^{\epsilon\mu} + \Sigma_\epsilon \{\nu\epsilon, \mu\} A^{\lambda\epsilon},$$

$$(21) \qquad A_{\lambda\nu}^\mu = \frac{\partial A_\lambda^\mu}{\partial x_\nu} - \Sigma_\epsilon \{\nu\lambda, \epsilon\} A_\epsilon^\mu + \Sigma_\epsilon \{\nu\epsilon, \mu\} A_\lambda^\epsilon.$$

In a similar way one introduces also the covariant derivative $A_{\lambda\mu\nu}$ of $A_{\lambda\mu}$ by means of the definition

$$(22) \qquad A_{\lambda\mu\nu} = \frac{\partial A_{\lambda\mu}}{\partial x_\nu} - \Sigma_\epsilon \{\lambda\nu, \epsilon\} A_{\epsilon\mu} - \Sigma_\epsilon \{\mu\nu, \epsilon\} A_{\lambda\epsilon}.$$

In each case it may be shown that the covariant derivative has the character indicated by the notation adopted for it.

Whenever the Christoffel symbols vanish the covariant derivatives reduce to ordinary derivatives; they so reduce, in particular, when the g's have Galilean or any other constant values.

§ 43. The Riemann-Christoffel Tensor.

Let A_μ be any covariant tensor of rank unity. Form its covariant derivative $A_{\mu\nu}$ in accordance with equation (18). Form the covariant derivative $A_{\mu\nu\sigma}$ of $A_{\mu\nu}$ in accordance with equation (22), employing the form of $A_{\mu\nu}$ given in (18). Thus we have

$$A_{\mu\nu\sigma} = \frac{\partial^2 A_\mu}{\partial x_\sigma \partial x_\nu} - \Sigma_\rho \{\mu\nu, \rho\} \frac{\partial A_\rho}{\partial x_\sigma} - \Sigma_\epsilon \{\mu\sigma, \epsilon\} \frac{\partial A_\epsilon}{\partial x_\nu} - \Sigma_\epsilon \} \nu\sigma, \epsilon\} \frac{\partial A_\mu}{\partial x_\epsilon}$$

$$+ \Sigma_{\epsilon\rho} \{\nu\sigma, \epsilon\} \{\mu\epsilon, \rho\} A_\rho + \Sigma_{\epsilon\rho} \{\mu\sigma, \epsilon\} \{\epsilon\nu, \rho\} A_\rho - \Sigma_\rho A_\rho \frac{\partial}{\partial x_\sigma} \{\mu\nu, \rho\}.$$

In view of the definitions of the Christoffel symbols it is seen that the first five terms in the second member are unaltered

by an interchange of ν and σ. Hence the tensor $A_{\mu\nu\sigma}-A_{\mu\sigma\nu}$ has the value $\Sigma_\rho B^\rho_{\mu\nu\sigma}A_\rho$ where

$$(23)\qquad B^\rho_{\mu\nu\sigma}=\Sigma_\epsilon\{\mu\sigma,\ \epsilon\}\{\epsilon\nu,\ \rho\}-\Sigma_\epsilon\{\mu\nu,\ \epsilon\}\{\epsilon\sigma,\ \rho\}$$
$$+\frac{\partial}{\partial x_\nu}\{\mu\sigma,\ \rho\}-\frac{\partial}{\partial x_\sigma}\{\mu\nu,\ \rho\}.$$

Since A_ρ is arbitrary and $\Sigma_\rho B^\rho_{\mu\nu\sigma}A_\rho$ has the tensor character $B_{\mu\sigma\nu}$ it follows that $B^\rho_{\mu\nu\sigma}$ has the tensor character anticipated by the notation employed. It is called the *Riemann-Christoffel tensor*. It will be observed that this tensor depends upon nothing but the fundamental tensor $g_{\mu\nu}$.

Now if we have a physical situation in which it is possible to choose the coordinate system so that the coefficients $g_{\mu\nu}$ shall be constants, then for this system the $B^\rho_{\mu\nu\sigma}$ must vanish. From this and its tensor character it follows that it must also vanish however the coordinate system is transformed in accordance with equations (9). The vanishing of this tensor is then a necessary condition that it shall be possible to choose the system of reference in such wise that the $g_{\mu\nu}$ shall have their constant Galilean values. It may be shown (though we do not here give the proof) that the condition is also sufficient.

From this it follows that in our problem the vanishing of the Riemann-Christoffel symbol corresponds to the possibility of choice of coordinates such that the special theory of relativity shall be valid in a finite region.

If we employ the notation

$$R_{\mu\nu}=-\Sigma_\alpha\frac{\partial}{\partial x_\alpha}\{\mu\nu,\ \alpha\}+\Sigma_{\alpha\beta}\{\mu\alpha,\ \beta\}\{\nu\beta,\ \alpha\},$$

$$S_{\mu\nu}=\frac{\partial^2\log\sqrt{-g}}{\partial x_\mu\partial x_\nu}-\Sigma_\alpha\{\mu\nu,\ \alpha\}\frac{\partial\log\sqrt{-g}}{\partial x_\alpha},$$

we have without difficulty the relation

$$(24)\qquad B_{\mu\nu}\equiv\Sigma_\tau B^\tau_{\mu\nu\tau}=R_{\mu\nu}+S_{\mu\nu}.$$

In order to effect the simplification in the result stated here one needs the following value of $\Sigma_\rho\{\mu\rho,\ \rho\}$:

$$\Sigma_\rho\{\mu\rho,\ \rho\} = \tfrac{1}{2}\Sigma_{\rho\epsilon} g^{\rho\epsilon}\left(\frac{\partial g_{\mu\epsilon}}{\partial x_\rho} + \frac{\partial g_{\rho\epsilon}}{\partial x_\mu} - \frac{\partial g_{\mu\rho}}{\partial x_\epsilon}\right)$$

$$= \tfrac{1}{2}\Sigma_{\rho\epsilon} g^{\rho\epsilon}\frac{\partial g_{\rho\epsilon}}{\partial x_\mu}$$

$$= \frac{1}{2g}\frac{\partial g}{\partial x_\mu} = \frac{\partial}{\partial x_\mu}\log\sqrt{-g},$$

the second last member being obtained by use of the fact that $g^{\rho\epsilon}g$ is the cofactor of $g_{\rho\epsilon}$ in the determinant g.

It may readily be shown that $R_{\mu\nu}$ and $S_{\mu\nu}$ have the tensor character indicated by the notation.

Now we have seen that the coordinate axes may be chosen so that $\sqrt{-g}$ has the value 1. Under such choice several of the preceding formulæ become simpler. This is particularly true of the expression for $B_{\mu\nu}$ since $S_{\mu\nu}$ then has the value zero. This simplification is of considerable importance in the theory of gravitation on account of the fundamental role in this theory of the tensor $B_{\mu\nu}$.

§ 44. Einstein's Law of Gravitation.

We have seen (in § 40) that the values of the g's which obtain in the absence of a gravitational field and for a suitable choice of reference system are the constant Galilean values. In § 43 we saw that the vanishing of the Riemann-Christoffel tensor $B^\rho_{\mu\nu\sigma}$ is a necessary and sufficient condition that it shall be possible to choose the system of reference in such wise that the $g_{\mu\nu}$ shall have the constant Galilean values. Hence a necessary and sufficient condition for the absence of a permanent gravitational field is the following:

$$(25)\qquad \Sigma_\epsilon\{\mu\sigma,\ \epsilon\}\{\epsilon\nu,\ \rho\} - \Sigma_\epsilon\{\mu\nu,\ \epsilon\}\{\epsilon\sigma,\ \rho\} + \frac{\partial}{\partial x_\nu}\{\mu\sigma,\ \rho\} - \frac{\partial}{\partial x_\sigma}\{\mu\nu,\ \rho\} = 0.$$

Of the 96 relations obtained from the six effectively distinct combinations of σ and ν and the 16 combinations of μ and ρ,

only 20 are independent. If we write $(\mu\tau\sigma\nu)$ for $\Sigma_\rho g_{\tau\rho} B^\rho_{\mu\nu\sigma}$ so that $B^\rho_{\mu\nu\sigma} = \Sigma_\lambda g^{\lambda\rho}(\mu\lambda\sigma\nu)$, it is seen that equation (25) is equivalent to the equation $(\mu\tau\sigma\nu) = 0$. The reduction to 20 independent relations is then effected through use of the identities

$$(\mu\tau\sigma\nu) \equiv -(\tau\mu\sigma\nu) \equiv (\nu\sigma\tau\mu) \equiv (\sigma\nu\mu\tau),$$
$$(\mu\tau\sigma\nu) + (\mu\sigma\nu\tau) + (\mu\nu\tau\sigma) \equiv 0.$$

Now the general law of gravitation must contain as a special case that expressed by the vanishing of the Riemann-Christoffel symbol and must itself be expressed in the form of differential equations satisfied by the g's. One of the simplest conditions meeting these requirements is that expressed by the vanishing of the tensor $B_{\mu\nu}$ defined in equations (24). This yields the equation

$$(26) \qquad B_{\mu\nu} \equiv -\Sigma_\rho \frac{\partial}{\partial x_\rho}\{\mu\nu, \rho\} + \Sigma_{\rho\epsilon}\{\mu\rho, \epsilon\}\{\nu\epsilon, \rho\} + \frac{\partial^2}{\partial x_\mu \partial x_\nu} \log \sqrt{-g} - \Sigma_\epsilon\{\mu\nu, \epsilon\}\frac{\partial}{\partial x_\epsilon} \log \sqrt{-g} = 0.$$

Since $B_{\mu\nu} = B_{\nu\mu}$ there are in (26) only ten different equations; and among these there are four identical relations, so that only six of the equations are independent. These equations are taken by Einstein in the general theory of relativity as the mathematical expression of the law of gravitation in the absence of matter and the electromagnetic field. [The reader should observe the marked simplification of the equations when the reference system is so chosen that $\sqrt{-g} = 1$.]

Einstein insists that there is a minimum of arbitrariness connected with this choice of equations. For $B_{\mu\nu}$ is the only tensor of second rank which is formed from the $g_{\mu\nu}$ and their first and second derivatives and is linear in the second derivatives. Moreover, no tensor of lower rank can be built up out of the components of $B^\rho_{\mu\nu\sigma}$ by allowable processes. One who counts up all the terms represented by the various symbols in (26) will probably admit that we should first find out whether the suggested law of gravitation is valid before trying a more

complicated one. But it must be remembered that the general theory, apart from facts of observation, does not lead necessarily to this particular law.

In the Newtonian theory of attraction it is the Laplace equation $\Delta^2\phi=0$ which corresponds to our equation (26). The covariant equation corresponding to Poisson's equation $\Delta^2\phi=-4\pi\rho$ has also been treated by several writers (including Einstein); but we shall not develop it here.

§ 45. The Motion of a Particle.

Denote by A^σ the contravariant vector $\partial x_\sigma/\partial s$. Multiplying A^α by the covariant derivative $A_\alpha{}^\sigma$ of A^σ,

$$A_\alpha{}^\sigma=\frac{\partial}{\partial x_\alpha}\left(\frac{\partial x_\sigma}{\partial s}\right)+\Sigma_\beta\{\alpha\beta,\ \sigma\}\frac{\partial x_\beta}{\partial s},$$

and summing as to α, we have

$$\Sigma_\alpha A^\alpha A_\alpha{}^\sigma=\frac{d^2x_\sigma}{ds^2}+\Sigma_{\alpha\beta}\{\alpha\beta,\ \sigma\}\frac{\partial x_\alpha}{ds}\frac{\partial x_\beta}{ds}.$$

From this equation it follows that the expression in the second member is a contravariant vector. Let us consider the equation obtained by setting it equal to zero, namely,

$$\text{(27)}\qquad \frac{d^2x_\sigma}{ds^2}+\Sigma_{\alpha\beta}\{\alpha\beta,\ \sigma\}\frac{\partial x_\alpha}{\partial s}\frac{\partial x_\beta}{\partial s}=0,\quad \sigma=1,\ 2,\ 3,\ 4.$$

On account of the covariant character of this equation it is satisfied or not independently of the choice of coordinates. In the case of the special theory of relativity the Christoffel symbols have the value zero and the equations reduce to $d^2x_\sigma/ds^2=0$; and these are equations of a straight line. In this special case, then, equations (27) give the path of a moving particle in the absence of a permanent gravitational field. Accordingly, in view of our principle of equivalence, we assume that equations (27) are the equations of motion of a particle referred to any axes, even though there is a permanent gravitational field. We are justified in this in view of the fact that

the equations contain no derivative of the $g_{\mu\nu}$ of order higher than the first.

Let us now see in what sense the Newtonian theory of the motion of a particle is a sort of first approximation to the theory based on equations (27). In the case of the special theory of relativity the components dx_1/ds, dx_2/ds, dx_3/ds of the velocity v can have arbitrary values. If the velocity of light is taken to be unity and v is a very small quantity then its components are small while dx_4/ds is equal to unity except for quantities of the second order. Moreover, in the limiting case of the special theory of relativity the quantities $g_{\mu\nu}$ have the value zero when $\mu \neq \nu$, while $g_{11} = -1$, $g_{22} = -1$, $g_{33} = -1$, $g_{44} = 1$. From this point of view the quantities $\{\alpha\beta, \sigma\}$ are of order at least as high as the first. From these results Einstein concludes that in equation (27) the desired approximation is to be attained by considering only that $g_{\mu\nu}$ for which $\mu = 4 = \nu$ and is then led to the approximate equations

$$\frac{d^2x_4}{dt^2} = -\{44, 4\}, \qquad \frac{d^2x_\sigma}{dt^2} = \{44, \sigma\}, \qquad \sigma = 1, 2, 3,$$

by taking the approximate relation $ds = dt$.

If we assume further that the gravitational field is quasi-static in the sense that the acceleration due to the gravitational field is very small compared with the velocity of light so that derivatives with respect to the time may be neglected in comparison with those with respect to the space-coordinates, we have approximately

$$\frac{d^2x_\sigma}{dt^2} = -\frac{1}{2}\frac{\partial g_{44}}{\partial x_\sigma}, \qquad \sigma = 1, 2, 3.$$

These are the equations of motion of a material particle in the Newtonian theory in which $g_{44}/2$ plays the role of the gravitational potential. It is noteworthy that these approximate equations depend on the single component g_{44} of the fundamental tensor.

Going back to the general point of view of equations (27)

and accepting the guidance of our current ideas (just seen to be approximately valid), let us consider the case of a particle at rest at the origin of coordinates in our space-time extension. We depart somewhat from the strict standpoint of the general theory of relativity and choose

$$(28) \qquad x_1=r, \quad x_2=\theta, \quad x_3=\phi, \quad x_4=t$$

as our coordinates, treating them as the usual polar coordinates. Then ds^2 may be assumed to have the form

$$ds^2=-e^{\lambda}dr^2-e^{\mu}(r^2d\theta^2+r^2\sin^2\theta\,d\phi^2)+e^{\nu}dt^2,$$

where λ, μ, ν are functions of r only.

One may justify the omission of the product terms $drd\theta$ and $drd\phi$ and $d\theta d\phi$ by the symmetry of the polar coordinates, and the omission of $drdt$ and $d\theta dt$ and $d\phi dt$ by the symmetry of a static field with respect to past and future time.

If we write $r^2e^{\mu}=r'^2$ and absorb into the λ the resulting change in dr^2 and then write r for the new r' we have for ds^2 the expression

$$(29) \qquad ds^2=-e^{\lambda}dr^2-r^2d\theta^2-r^2\sin^2\theta d\phi^2+e^{\nu}dt^2.$$

From this we have $g_{\mu\nu}=0$ when $\mu\neq\nu$ and

$$(30) \qquad g_{11}=e^{-\lambda}, \quad g_{22}=-r^2, \quad g_{33}=-r^2\sin^2\theta, \quad g_{44}=e^{\nu}.$$

Here λ and ν are functions of r.

The determinant g reduces to its main diagonal and we have

$$-g=e^{\lambda+\nu}r^4\sin^2\theta, \quad g^{\sigma\sigma}=1/g_{\sigma\sigma}.$$

The three-index symbol $\{\sigma\tau, \alpha\}$ has the value

$$\{\sigma\tau, \alpha\}=\frac{1}{2g_{\alpha\alpha}}\left\{\frac{\partial g_{\alpha\sigma}}{\partial x_\tau}+\frac{\partial g_{\alpha\tau}}{\partial x_\sigma}-\frac{\partial g_{\sigma\tau}}{\partial x_\alpha}\right\}.$$

If σ, τ, ρ are three distinct numbers we have $\{\sigma\tau, \rho\}=0$ while

$$(31)\quad \{\sigma\sigma,\sigma\}=\tfrac{1}{2}\frac{\partial}{\partial x_\sigma}\log g_{\sigma\sigma},\quad \{\sigma\sigma,\tau\}=-\frac{1}{2g_{\tau\tau}}\frac{\partial}{\partial x_\tau}g_{\sigma\sigma},$$

$$\{\sigma\tau,\tau\}=\tfrac{1}{2}\frac{\partial}{\partial x_\sigma}\log g_{\tau\tau}.$$

By a straightforward computation one may now evaluate the Christoffel symbols involved in (26) and so obtain the explicit form of those equations for the special case now in consideration. It turns out that $B_{\mu\nu}=0$ is identically satisfied when $\mu\neq\nu$. From the four equations in which $\mu=\nu$ we have by the indicated direct computation the relations

$$(32)\quad \begin{aligned} &\tfrac{1}{2}\nu''-\tfrac{1}{4}\lambda'\nu'+\tfrac{1}{4}\nu'^2-\lambda'/r=0,\\ &e^{-\lambda}[1+\tfrac{1}{2}r(\nu'-\lambda')]-1=0,\\ &\sin^2\theta\cdot e^{-\lambda}[1+\tfrac{1}{2}r(\nu'-\lambda')]-\sin^2\theta=0,\\ &e^{\nu-\lambda}(-\tfrac{1}{2}\nu''+\tfrac{1}{4}\lambda'\nu'-\tfrac{1}{4}\nu'^2-\nu'/r)=0,\end{aligned}$$

where the primes denote differentiation with respect to r.

Combining the first and last equations we see that $\lambda'=-\nu'$. Then, since λ and ν must tend to zero as r tends to infinity so that g_{11} and g_{44} shall have at infinity the Galilean values -1 and $+1$, respectively, it follows that $\lambda=-\nu$. Then the second and third equations in (32) both reduce to the equation

$$e^\nu(1+r\nu')=1.$$

Solving this we have $e^\nu=1-2m/r$, where $2m$ is a constant of integration. This solution also satisfies the first and fourth equations in (32). Substituting in (29) the derived values for e^λ and e^ν we have

$$(33)\quad ds^2=-\left(1-\frac{2m}{r}\right)^{-1}dr^2-r^2d\theta^2-r^2\sin^2\theta d\phi^2 +\left(1-\frac{2m}{r}\right)dt^2.$$

Substituting into (27) the values of the Christoffel symbols in the same special forms as we have just employed in (26) we

may obtain the explicit forms of equations (27) for our present problem. Equation (27) for $\sigma=2$ thus becomes

$$\text{(34)}\qquad \frac{d^2\theta}{ds^2}-\cos\theta\sin\theta\left(\frac{d\phi}{ds}\right)^2+\frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds}=0.$$

If we choose coordinates so that the particle moves initially in the plane $\theta=\pi/2$ we have initially $d\theta/ds=0$ and $\cos\theta=0$, so that $d^2\theta/ds^2=0$; whence it follows that the particle continues to move in this plane. With such a choice of coordinates equations (27) for $\sigma=1, 3, 4$ take the forms

$$\text{(35)}\qquad \frac{d^2r}{ds^2}+\tfrac{1}{2}\frac{d\lambda}{dr}\left(\frac{dr}{ds}\right)^2-re^{-\lambda}\left(\frac{d\phi}{ds}\right)^2+\tfrac{1}{2}e^{\nu-\lambda}\frac{d\nu}{dr}\left(\frac{dt}{ds}\right)^2=0,$$

$$\text{(36)}\qquad \frac{d^2\phi}{ds^2}+\frac{2}{r}\frac{dr}{ds}\frac{d\phi}{ds}=0,\qquad \frac{d^2t}{ds^2}+\frac{d\nu}{ds}\frac{dt}{ds}=0.$$

From equations (36) we have

$$\text{(37)}\qquad r^2\frac{d\phi}{ds}=h,\qquad \frac{dt}{ds}=\gamma e^{-\nu}=\gamma\left(1-\frac{2m}{r}\right)^{-1},$$

where h and γ are constants of integration and where in the last member we have replaced $e^{-\nu}$ by its value obtained in the paragraph ending with equation (33). Again, if we replace λ and ν in (35) by their values previously derived we have

$$\text{(38)}\qquad \frac{d^2r}{ds^2}-\frac{m}{r^2}\left(1-\frac{2m}{r}\right)^{-1}\left(\frac{dr}{ds}\right)^2-r\left(1-\frac{2m}{r}\right)\left(\frac{d\phi}{ds}\right)^2$$
$$+\frac{m}{r^2}\left(1-\frac{2m}{r}\right)\left(\frac{dt}{ds}\right)^2=0.$$

Since $d\theta=0$ and $\theta=\pi/2$ we now have from (33) the relation

$$\text{(39)}\qquad \left(1-\frac{2m}{r}\right)^{-1}\left(\frac{dr}{ds}\right)^2+r^2\left(\frac{d\phi}{ds}\right)^2-\left(1-\frac{2m}{r}\right)\left(\frac{dt}{ds}\right)^2+1=0.$$

From this relation and (37) we have

$$\text{(40)}\qquad \left(\frac{dr}{ds}\right)^2+r^2\left(\frac{d\phi}{ds}\right)^2=\gamma^2-1+\frac{2m}{r}+2m\frac{h^2}{r^3}.$$

If in equation (40) we replace $d\phi/ds$ by its value from (37) and then differentiate with respect to s we have for d^2r/ds^2 the same value as that obtained from (38) on eliminating the first derivatives in (38) by aid of (37) and (39). Hence, if we retain equations (37) and (40) we may omit equation (38). Then equations (37) and (40) are the sole equations of motion.

In the corresponding coordinates the Newtonian equations of elliptic motion are

$$(41) \qquad \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\phi}{dt}\right)^2 = -\frac{m}{a} + \frac{2m}{r}, \qquad r^2\frac{d\phi}{dt} = h.$$

To make the first of these correspond with (40) we must regard ds as replacing dt and take for γ^2 the value $\gamma^2 = 1 - m/a$, a being the semimajor axis of the orbit. The term $2mh^2/r^3$ in (40) represents a small additional effect not in evidence in the Newtonian theory. The quantity m, previously introduced as a constant of integration, is now to be identified as the mass of the attracting particle measured in gravitational units.

§ 46. Three Crucial Phenomena.

If we take one kilometre as the unit of length and choose the unit of time so that the velocity of light is unity, we obtain from (41) the approximate value $m = 1.47$ for the mass of the sun on supposing that the path of motion of the earth is circular so that $r = a = 1 \cdot 49.10^8$ and on employing the value $\omega = 6 \cdot 64.10^{-13}$ of the angular velocity $d\phi/dt$. Hence m/r is of the order 10^{-8}. Moreover, it follows from the second equation in (41) that h^2/r^2 is approximately equal to $\omega^2 a^2$ and is hence of order 10^{-8}.

From (40) and the first equation in (37) we have

$$\left(\frac{h}{r^2}\frac{dr}{d\phi}\right)^2 + \frac{h^2}{r^2} = \gamma^2 - 1 + \frac{2m}{r} + 2m\frac{h^2}{r^3}.$$

If we transform this equation by use of the relation $r = 1/u$ and differentiate both members of the resulting equation with respect to ϕ, we have

$$\frac{d^2u}{d\phi^2}+u=\frac{m}{h^2}+3mu^2. \tag{42}$$

Now h^2u^2 is a quantity of the order of 10^{-8}; hence, we may get a roughly approximate solution of (42) by neglecting the term $3mu^2$. This gives

$$u=\frac{m}{h^2}[1+e\cos(\phi-w)], \tag{43}$$

where e and w are constants of integration. In order to get a second approximation to the solution we substitute the value of u in the last term of (42) and obtain the equation

$$\frac{d^2u}{d\phi^2}+u=\frac{m}{h^2}+\frac{6m^3}{h^4}e\cos(\phi-w)+\frac{3m^3}{h^4}+\frac{3m^3e^2}{2h^4}[1+\cos 2(\phi-w)].$$

Through use of the approximate value of u in (43) and the fact that h^2u^2 is of order 10^{-8}, it may be seen that the third and fourth terms in the second member of the last equation cannot produce appreciable effects. But the second term is of a suitable period to produce an increasing effect by resonance. Retaining from the second member only the first and second terms and solving the equation so curtailed we have

$$u=\frac{m}{h^2}[1+e\cos(\phi-w)+\frac{3m^2}{h^2}\phi e\sin(\phi-w)].$$

Writing δw for $3m^2\phi/h^2$ and neglecting the second and higher powers of δw, we may put the approximate value of u in the form

$$u=\frac{m}{h^2}[1+e\cos(\phi-w-\delta w)].$$

Applying this to the case of a planet moving around the sun we find that while the planet moves through one revolution the perihelion advances by a fraction of a revolution equal to

$$\frac{\delta w}{\phi}=\frac{3m^2}{h^2}=\frac{3m}{a(1-e^2)}=\frac{12\pi^2a^2}{c^2T^2(1-e^2)},$$

T being the period of the planet and c being the velocity of light in customary units introduced into the last member for convenience.

This formula gives 42.9, 8.6, 3.8, 1.35 seconds for the respective advances of the perihelion (per century) of the four inner planets, Mercury, Venus, Earth, Mars. This is in close agreement with observation. Thus the theory of Einstein yields a very satisfactory explanation of the celebrated large discordance in the motion of the perihelion of Mercury which has occupied the attention of astronomers since the time of Leverrier. There is no trace of forced agreement in connection with this remarkable success of the theory.

A second crucial phenomenon for the theory is that of the deflection of a ray of light. In the absence of a gravitational field and with the choice of coordinates which we have been employing the velocity of light has the constant value unity. Hence

$$\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2=1,$$

so that

$$ds^2=-dx^2-dy^2-dz^2+dt^2=0.$$

Therefore, for the motion of light in the absence of a gravitational field we have $ds=0$; then by the principle of equivalence we must also have $ds=0$ even in a gravitational field. Employing this value of ds and assuming that the path of light is in the plane $\theta=\pi/2$ we have from (33) the relation

$$\left(1-\frac{2m}{r}\right)^{-1}\left(\frac{dr}{dt}\right)^2+\left(r\frac{d\phi}{dt}\right)^2=1-\frac{2m}{r}. \tag{44}$$

If v is the velocity of light in a direction making an angle α with the radius vector we have

$$v^2\left\{\left(1-\frac{2m}{r}\right)^{-1}\cos^2\alpha+\sin^2\alpha\right\}=1-\frac{2m}{r},$$

whence

$$v=\left(1-\frac{2m}{r}\right)\left\{1-\frac{2m}{r}\sin^2\alpha\right\}^{-\frac{1}{2}}.$$

Since this gives a velocity for light varying with the direction we alter our coordinates slightly by replacing r by $r+m$, whence we have that r^2 is to be replaced by a quantity approximately equal to

$$r^2\left(1-\frac{2m}{r}\right)^{-1}.$$

With such a value of r equation (44) becomes approximately the equation

$$\left(\frac{dr}{dt}\right)^2+\left(r\frac{d\phi}{dt}\right)^2=\left(1-\frac{2m}{r}\right)^2,$$

whence we obtain for v the approximate value

$$v=1-\frac{2m}{r},$$

the same in all directions.

If we employ the principle that the course of a ray of light depends only on the variation of velocity, we find that it will be the same as in a Euclidean space filled with material of a refractive index μ given by the relation $\mu=1/v$, and hence approximately by the relation

$$\mu=1+\frac{2m}{r}.$$

The gravitational field round a particle thus acts as a converging lens.

From the foregoing value of μ it may be shown without difficulty that a ray of light from $-\infty$ to $+\infty$, which passes at a distance R from a particle of mass m, will experience a total deflection of amount $4m/R$. For the sun we have $m=1.47$ and $R=$sun's radius$=697{,}000$ km. Hence for a star seen close to the limb of the sun we shall have a deflection of 1.74 seconds of angular measure.

This prediction was tested by observations made independently at two stations during the eclipse of the sun of May 29, 1919; the values for the deflection obtained at the two stations

are 1.61 and 1.98 seconds of angular measure, results in substantial agreement with the predicted value.

Thus the Einstein law of gravitation, as expressed in equation (33), has been checked for high velocities by the deflection of a ray of light and for comparatively low velocities by the motion of the perihelion of Mercury—two very remarkable conquests to be made simultaneously by a single theory.

A third crucial phenomenon is afforded by the vibration of an atom in a gravitational field. Such an atom is a natural clock and should therefore give an invariant measure of an interval of time. If the atom is at rest in the system of coordinates (which themselves may be in motion) we have $dx=dy=dz=0$, so that $ds^2=g_{44}dt^2$. If we have two similar atoms at different parts of the field where the potentials are g_{44} and g'_{44}, respectively, we have from the invariance of ds that

$$\sqrt{g_{44}}dt=\sqrt{g'_{44}}dt'.$$

If t refers to the photosphere of the sun where $g_{44}=1-2m/R$, R being the sun's radius, and t' refers to a point on the earth where g'_{44} is sensibly equal to unity, we have approximately,

$$\frac{dt}{dt'}=1+\frac{m}{R}=1.00000212.$$

From this it follows that the atom vibrates more slowly on the sun than on the earth, and hence that the lines of the spectrum should be displaced toward the red. For the part of the spectrum usually observed this displacement amounts to about .008 tenth-meters (a tenth-meter $=10^{-10}$ meters).

There is not yet agreement as to whether the phenomenon thus predicted is actually existent; in fact, some doubt has been felt as to whether the argument is valid by which this prediction is supported. C. E. St. John (*Astrophysical Journal*, vol. 46, p. 249) has given negative evidence and L. Grebe and A. Bachem (*Deut. Phys. Gesell., Verh.*, vol. 21, p. 454) have given positive evidence for the existence of the effect. Einstein (quoted in *Science* for March 12, 1920, p. 270) seems to believe

that the existence of the phenomenon has been established. But, if it should turn out that the test fails, the most appropriate conclusion would seem to be that we have insufficient knowledge of the conditions of atomic vibrations rather than that the theory of relativity is thus discredited.

§ 47. The Electromagnetic Equations.

That Maxwell's equations may be reduced to a covariant form and hence that all electromagnetic phenomena described by them are in agreement with the principle of relativity may readily be shown in the following way (the exposition being based on that of Eddington, l.c., pp. 76–77):

The electromagnetic field is described by a covariant vector κ_μ which in Galilean coordinates has the components

$$\kappa_\mu = (-F,\ -G,\ -H,\ \Phi), \tag{45}$$

where F, G, H is the vector potential and Φ is the scalar potential of the ordinary theory. If $\kappa_{\mu\nu}$ is the covariant derivative of κ_μ we have by (18)

$$\frac{\partial \kappa_\mu}{\partial x_\nu} - \frac{\partial \kappa_\nu}{\partial x_\mu} = \kappa_{\mu\nu} - \kappa_{\nu\mu}, \tag{46}$$

a covariant tensor which we denote by $F_{\mu\nu}$.

The electric and magnetic forces of the usual theory are expressed in our present notation by formulas like

$$X = \frac{\partial \kappa_1}{\partial x_4} - \frac{\partial \kappa_4}{\partial x_1}, \qquad \alpha = \frac{\partial \kappa_2}{\partial x_3} - \frac{\partial \kappa_3}{\partial x_2}. \tag{47}$$

Hence in Galilean coordinates the value of $F_{\mu\nu}$ is given by the array

$$\begin{array}{cccc} 0 & -\gamma & \beta & -X \\ \gamma & 0 & -\alpha & -Y \\ -\beta & \alpha & 0 & -Z \\ X & Y & Z & 0 \end{array}$$

where μ varies through a row and ν varies through a column.

The associated contravariant tensor $F^{\mu\nu}=\Sigma_{\alpha\beta}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}$ is given similarly by the array

$$\begin{array}{cccc} 0 & -\gamma & \beta & X \\ \gamma & 0 & -\alpha & Y \\ -\beta & \alpha & 0 & Z \\ -X & -Y & -Z & 0 \end{array}$$

In the ordinary theory the Maxwell equations may be written in the form

$$(48)\qquad \frac{\partial Z}{\partial y}-\frac{\partial Y}{\partial z}=-\frac{\partial \alpha}{\partial t},\quad \frac{\partial X}{\partial z}-\frac{\partial Z}{\partial x}=-\frac{\partial \beta}{\partial t},\quad \frac{\partial Y}{\partial x}-\frac{\partial X}{\partial y}=-\frac{\partial \gamma}{\partial t},$$

$$(49)\qquad \frac{\partial \gamma}{\partial y}-\frac{\partial \beta}{\partial z}=\frac{\partial X}{\partial t}+u,\quad \frac{\partial \alpha}{\partial z}-\frac{\partial \gamma}{\partial x}=\frac{\partial Y}{\partial t}+v,\quad \frac{\partial \beta}{\partial x}-\frac{\partial \alpha}{\partial y}=\frac{\partial z}{\partial t}+w,$$

$$(50)\qquad \frac{\partial X}{\partial x}+\frac{\partial Y}{\partial y}+\frac{\partial Z}{\partial z}=\rho,$$

$$(51)\qquad \frac{\partial \alpha}{\partial x}+\frac{\partial \beta}{\partial y}+\frac{\partial \gamma}{\partial z}=0,$$

where the velocity of light is taken to be unity and the Heaviside-Lorentz unit of charge is chosen so that the factor 4π is absent.

The electric current u, v, w and density ρ form a contravariant vector J^μ, since

$$J^\mu\equiv(u,\ v,\ w,\ \rho)=\Sigma e\left(\frac{dx}{ds},\ \frac{dy}{ds},\ \frac{dz}{ds},\ \frac{dt}{ds}\right)$$

per unit volume. Equations (49) and (50) yield the relations

$$(52)\qquad \Sigma_\nu\frac{\partial F^{\mu\nu}}{\partial x_\nu}=J^\mu,$$

while equations (48) and (51) may be written

$$\frac{\partial F_{\mu\nu}}{\partial x_\sigma}+\frac{\partial F_{\nu\sigma}}{\partial x_\mu}+\frac{\partial F_{\sigma\mu}}{\partial x_\nu}=0.$$

From the definition of $F_{\mu\nu}$ it is seen that the last equation is satisfied identically, so that (46) and (52) represent the funda-

mental electromagnetic equations. The former is covariant and the latter is rendered so on replacing the ordinary derivative by the covariant derivative. Hence the required covariant equations take the form

$$F_{\mu\nu} = \frac{\partial \kappa_\mu}{\partial x_\nu} - \frac{\partial \kappa_\nu}{\partial x_\mu}, \qquad \Sigma_\nu F_\nu^{\mu\nu} = J^\mu.$$

These hold in the gravitational field because the conditions for the application of the principle of equivalence are satisfied.

Here we have shown that the electromagnetic equations are consistent with the theory of relativity, but we have not derived them by means of that theory. A more far-reaching treatment of the electromagnetic problem has been made by Weyl by means of a generalized form of the theory of relativity (*Annalen der Physik*, vol. 59 (1919), p. 101).

§ 48. Some General Considerations Relating to the Theory.

Silberstein (*Philosophical Magazine*, vol. 36 (1918), pp. 94–128) has undertaken to develop the general theory of relativity without the equivalence hypothesis. He concedes that the Einstein theory has one very strong point, namely, the requirement of general covariance of all physical laws, that actual phenomenal contents should be expressed (or at least be expressible) in a way showing their independence of the particular language or scaffolding adopted; but he insists that it has also a weak point, namely, the equivalence hypothesis which places gravitation, he believes, on an entirely exceptional and privileged footing, bringing it into intimate connection with the fundamental tensor which appears in the line-element of the world.

He proposes to retain the strong point and to reject the weak one and thus to develop the implications of the general principle of relativity without the equivalence hypothesis, in fact, without privileging gravitation at all. He considers the equivalence hypothesis to be a vulnerable point, independently

of agreement or disagreement with experimental facts, because of its special nature and of the great number of assumptions which it tacitly implies.

It is a matter of importance to separate these two elements and to ascertain to what extent the results obtained in the theory are based on the one or the other of the two parts of the general theory. In our treatment in the foregoing pages we have followed the method of Einstein and have not undertaken to separate the two elements or to distinguish between their consequences.

One matter not mentioned in our preceding treatment should have at least a word of attention. It will be observed that in the theory as developed the notion of gravitational force has hardly been present at all and that in fact the properties of the gravitational field are essentially geometrical rather than dynamical in character. Weyl, in the paper referred to at the end of § 47, has succeeded in extending the theory so as to include electromagnetic and gravitational forces in one geometrical scheme, thus extending the range in which the explanations may be stated in purely geometric terms.

When one undertakes to pursue to their extreme reach the geometric conceptions which thus arise one is soon brought to consider the fundamental character of the four-dimensional space-time extension by means of which the phenomena are thus interpreted geometrically. From this point of view one seems forced to conclude that our space-time extension is not a flat space like a plane or a Euclidean space of three dimensions, but has an essential curvature involved in its four-dimensional continuum analogous to that of a two-dimensional continuum represented by a warped surface in our usual space of three dimensions.

Some of those who have followed up this idea have come to the conclusion that our actual space-time manifold is finite in extent.

Several writers have considered the possibility of founding the entire theory of relativity on a certain different basis from

that employed in this chapter, namely, on the principle of least action. This and the topic just mentioned previously are treated in the last two chapters of the monograph of Eddington already referred to. For our purpose it suffices to say that it is possible to formulate the principle of least action naturally in such a way that our basic equations are equivalent to the principle. From a theoretical point of view there is much to be said in favor of developing the theory in this way; but the purposes of an elementary exposition are better served by the plan of treatment which we have adopted.

Whatever may be the final verdict as to the validity of the theory of relativity as a whole, it seems practically certain that it has already made a fundamental and permanent contribution to astronomy in developing the modification of Newton's law of gravitation associated with equation (33), the new form of the law now having been checked experimentally in two very different ways and thus established on a particularly secure experimental basis. Two such conquests as those recorded in § 46 have seldom been made so nearly simultaneously by a single theory developed from one point of view consistently maintained throughout.

INDEX.